FACULTEIT ECONOMIE EN
BEDRIJFSWETENSCHAPPEN

Data-driven operations management at a large university hospital

Proefschrift voorgedragen tot
het behalen van de graad van
Doctor in de Toegepaste
Economische Wetenschappen

door

Michael SAMUDRA

Number 552 2016

Doctoral Committee

Advisor:	Prof. dr. Erik Demeulemeester KU Leuven
Co-advisor:	Prof. dr. Brecht Cardoen Vlerick Business school, KU Leuven
Members:	Prof. dr. Inneke Van Nieuwenhuyse KU Leuven
	Prof. dr. Frank Rademakers University Hospital Leuven
	Prof. dr. Sally Brailsford University of Southampton

Acknowledgments

I have been lucky to have a great deal of help over the entire course of this PhD. Firstly, I would like to express my sincere gratitude to my supervisor Erik Demeulemeester. One of the reasons why Erik is a great supervisor is his ability to adapt to the style of the PhD student. At the start of my PhD, he supported me in finding a problem setting that is primarily motivated by practical relevance, which is an aspect that is really important to me in research. Next to my time in Leuven, I was fortunate to enjoy many conferences together with Erik. These were not only inspiring from a professional point of view, but also very pleasant from a personal point of view. Moreover, it was incredible how Erik always did the effort to make time when you asked for it, even when he was extremely busy, something I highly appreciated in our collaboration. This was a true asset as a short meeting with him was often enough to get out of a deadlock.

For continuous advice and support, I am also indebted to my co-supervisor Brecht Cardoen. His critical feedback has been to my great advantage. Working on a project for a longer time, lets one forget to question the validity of initially made assumptions. Brecht is the person who thinks out of the box and challenges those assumptions. He helped me to become more confident in the validity of the results presented in this thesis. Subsequently, I would also like to thank the other members of the committee. I would like to express my gratitude to Inneke Van Niewenhuyse for her detailed feedback. Her questions on how different surgical disciplines perform individually enabled me to greatly improve this part of the thesis. Additionally, I thank her for being the irreplaceable and caring women she is on our floor. Additionally, many thanks to Sally Brailsford for her valuable feedback. I truly appreciate her willingness to travel to Leuven at the most inconvenient times. It was always a pleasure to meet her on conferences and to listen to her dedication and passion for healthcare research.

My sincere thanks also go to Frank Rademakers. His deep understanding of the hospital setting and its processes together with his interest and quick grasp of operations management methods made him an irreplaceable element in the co-operation with the hospital. I also thank the rest of the team of the University Hospital Leuven, who provided me with the opportunity to work on a real setting and who gave us access to remarkably detailed and high quality data. Special thanks to Nancy Vansteenkiste, Pierre Luysmans, Christian Lamote, Philip Monnens, Jo Vandersmissen and Guido De Voldere. Without their precious support it would not have been possible to conduct this research. The discussions at the hospital were always inspiring, to the point and encouraging. I admire the hospital team's striving to continuously improve their operations and their openness to new ideas. It was an honor and a pleasure to work together with such a center of excellence.

Institutionally, I also have debts. It has been an amazing experience to be a researcher at the KU Leuven in this great city. There are few places where one can meet this easily people with such interesting and diverse backgrounds. The academic diversity and the various facilities provided by this university in a city that is great to live in make for a truly astonishing student experience.

As with any major piece of research and writing, there are many people who indirectly contributed in the completion of this thesis. For this reason, I would like to thank all colleagues of my research group for providing a fun and pleasant work atmosphere. I would like to thank: Ann, Catherine, Dennis, Gert, Guoxuan, Hamed, Jiangjiang, Jeroen, Joeri, Joren, Jorne, Kris, Marc, Maud, Mieke, Morteza, Nico, Nishant, Philippe, Pieter, Raïsa, Robert, Sarah, Shirley, Stef, Stefan, Valeria, Willy and Yannick. Thanks for the BBQs and cake Mondays and for the relaxing coffee breaks. Special thanks to the old gang for the vibrant office life at the start of my PhD. I would also like to thank the 5th floor for the valuable conversations we had and I am convinced that your conviction for mathematical proofs and your pursuit of trying to convince our floor of their importance makes all our research better. Therefore thanks to: Annette, the two

Barts, Dries, Fabrice, Kris, Vikram, Ward and Wenchao. My thanks also go to the MSI group. After all this time, I still do not know much about the research you do (I did learn that you prefer managerial insights over numbers) but I really enjoy our activities after office hours. Special thanks to Philippe and Manuel for all the good times.

Thanks also to my closest friends: Tibi, Dávid and Gábor. I really appreciate that even in times when meeting up on a regular basis became more difficult, everything remained like in old times. My thanks also go to Bori and Adi, I always enjoy the nice activities we do together and really value our friendship.

At the end comes the most personal of gratitude. Special thanks to my parents, who can say little more about this thesis than that it is about surgeries, but who mean the world to me. Thanks to my siblings, I am incredibly proud of you and I am really looking forward to see what you will become once you grow up. Finally, thanks to you Carla for being a partner in life and in science. I appreciate your continuous support and interest in my work. I cannot imagine how some of the chapters would have turned out without the countless discussions we had often until long into the night.

Michael Samudra

Abstract

In many hospitals there are patients who receive surgery later than what is medically indicated. In one of Europe's largest hospitals, the University Hospital Leuven, this is the case for approximately every third patient. Serving patients late cannot always be avoided as a highly utilized operating room (OR) department will sometimes suffer capacity shortage, occasionally leading to unavoidable delays in patient care. Nevertheless, serving patients late is a problem as it exposes them to an increased health risk and hence should be avoided whenever possible.

In order to improve the current situation, the delay in patient scheduling had to be quantified and the responsible mechanism, the scheduling process, had to be better understood. Drawing from this understanding, we implemented and tested realistic patient scheduling methods in a discrete-event simulation model.

In this text we describe some of the primary aspects and properties of the hospital's inpatient population, introduce the way patients are scheduled in reality and describe some of the major mechanisms that take place in the OR department. We will therefore describe patient arrival patterns, the relationship between estimated and realized surgery durations, the applied rescheduling mechanisms on the day of surgery and the non-elective to OR allocation schema. Finally, we will introduce some of the manually applicable scheduling methods and show how they perform in the resulting simulation environment.

We found that it is important to model non-elective arrivals and to include elective rescheduling decisions made on the surgery day itself. Rescheduling results in the fact that OR-related performance measures, such as overtime, will only loosely depend on the method with which patients are scheduled to a date and an OR before the day of the surgery.

Our results suggest that in case patients are scheduled to a final surgery date during their consultation session (one-step procedure), capacity considerations should guide the patient scheduling procedure because this will result in a high number of patients served within their medically indicated time limit (due time). Some of the tested methods that were expected to increase the percentage of high urgency patients being served within their due time (by reserving capacity for them) surprisingly did not work to that effect. This was due to the sporadic waste of the reserved capacity that, in the long run, leads to a decreased average capacity for all patient categories. One strategy that efficiently uses OR capacity is first come, first served. As applying first come, first served might not always be possible in a real setting, we found that it is important to allow for patient replanning.

Our results also suggest that in case patients are scheduled first to a week and, in a second step on a later date, to an exact weekday and OR (two-step procedure), it is very important that the second step is guided by the patient's urgency category. Additionally, it is important to allow higher urgency patients to be inserted at consultation time into the fixed weekly schedules created during the second step. Interestingly, we found that reserving a constant amount of capacity for high urgency patients is not necessarily beneficial from a whole system perspective.

As operations research results are most valuable if they are applied in practice, we also looked at three common pitfalls that could hamper the adoption of research results by stakeholders in this field: the lack of a clear choice of authors on whether to target researchers (contributing advanced methods) or practitioners (providing managerial insights), the use of ill-fitted performance measures in models and the failure to understandably report on the hospital setting and the method-related assumptions. We provide specific guidelines that help to avoid these pitfalls. First, we show how to build up an article based on the choice of the target group (i.e., researchers or practitioners). Making a clear distinction be-

tween target groups impacts the problem setting, the research task, the reported findings, and the conclusions. Second, we discuss points that need to be considered by researchers when deciding on the used performance measures. Third, we list the assumptions that need to be included in articles in order to enable readers to decide whether the presented research is relevant to them.

Contents

Chapter 1

Introduction

It is a problem if patients wait longer for surgery than what is deemed to be optimal by their surgeons. In those cases, patients are said to have been served after the due time (DT) [287], which can pose a health risk. In one of Europe's largest hospitals, the University Hospital Leuven in Belgium, 34.6% of patients are served after their target DT. This is not uncommon as a highly utilized OR department will sometimes suffer from a capacity shortage, occasionally leading to unavoidable delays in patient care.

Nevertheless, serving patients late should be prevented if possible, primarily from a medical standpoint, but also from a societal hidden cost perspective as patients in a worsened health condition are likely to require more resources.

In order to improve the current situation, the lateness of patients had to be quantified and the primarily responsible mechanism, which is the patient scheduling process, had to be better understood. Drawing from this understanding, we implemented and tested realistic patient scheduling policies using a discrete-event simulation model (DES). The results of the tests should help surgeons and nurses to better understand the consequences of their patient scheduling-related decisions.

The amount of time a patient can wait for surgery varies largely from case to

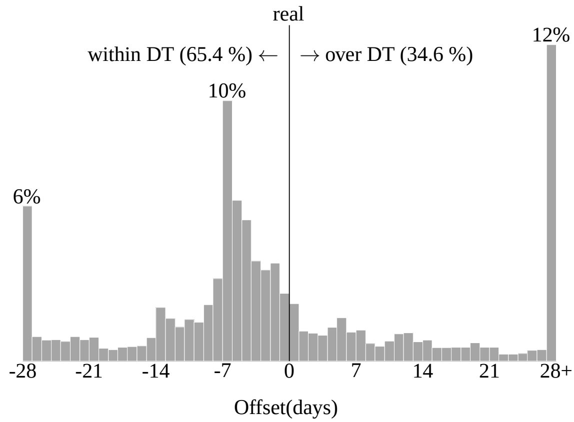


Fig. 1.1 The distribution of the number of patients served before/after their DT shows that most of them are served just before their DT. The open-ended histogram does not cover those electives that have not been assigned a DT and thus do not need to be served within a time limit. Including them and assuming they are always served within their DT, the total percentage of patients served within DT changes to 76%. The figure is based on data covering the entire years 2012-2013.

case. It depends on many factors such as the general health condition of the patient, the speed at which the underlying disease is progressing, the endured pain level and the detrimental lifestyle effects.

One way to ensure that patients receive surgery within an acceptable time limit is to enforce waiting time targets, such as DTs. DTs can be set up by the authority of a larger geographic region such as a government (e.g., Australia and Canada [6, 19]) or by a lower level authority such as a hospital.

DTs were set up at the University Hospital Leuven by the surgeons of the hospital and were determined on the basis of medical criteria. The DT is therefore a concept that has been implicitly considered, but has only been formalized recently. Formalizing it allows the hospital to use it as a benchmark criterion. Figure 1.1 shows that a large part of the patient population is served before their DT and around one third of them is served after their DT. The figure is based on data covering the entire years 2012-2013 including all 13 disciplines (Table 1.1) that are served in the hospital's 22 inpatient ORs.

The DT is assigned to patients by the respective physician in charge. It is divided

Table 1.1 There are thirteen disciplines served in the inpatient department.

GYN	Gynecology and obstetrics
Tx	Abdominal transplant surgery
ABD	Abdominal surgery
CAH	Cardiac surgery
NCH	Neurosurgery
ONC	General medical oncological
RHK	Plastic, reconstructive and cosmetic surgery
THO	Thoracic surgery
TRH	Traumatology
URO	Urology
VAT	Vascular surgery
MKA	Oral and maxillofacial surgery
NKO	Head and neck surgery

Table 1.2 The University Hospital Leuven uses eight DT categories.

	Category	Target time
Non-elective	1	Instantly
	2	Up to 6 hours
	3	Today
Elective	4	1 week
	5	1 - 2 weeks
	6	2 - 4 weeks
	7	4 - 8 weeks
	8	No target time

into eight categories (Table 1.2) where categories 4 to 8 are used to classify electives and categories 1 to 3 are used to classify non-electives. The DT of elective patients is defined in weeks whereas the DT of non-electives, as they have to be served the latest within 24 hours after their admittance, is defined in hours.

As even the least urgent non-elective patients have to be served within 24 hours, there is no room for scheduling them in advance. Non-electives are therefore not planned and they are only included into the simulation model to test their impact on the execution of the elective schedule.

As Table 1.2 shows, the DT is defined as a time interval suggesting that it is best for a patient to get surgery only after a certain reference period (DT start). It might seem unreasonable to let patients wait unnecessarily, but it can be the case that they or their surgeons need time to prepare for the surgery. For example, in a Dutch breast cancer center, patients would generally need 1 week to prepare for the intervention and undergo it within 5 weeks. From a scheduling perspective, the end time of the interval (DT end) is the determining factor.

The DT score of a discipline is calculated based on the weights associated with each DT category. The weights for DT categories 4 to 7 are 1, 1/2, 1/4 and 1/8 respectively. A weight of 0 is associated with DT category 8. The DT score of a discipline is the average DT weight assigned to their patients.

Figure 1.2 shows that both the DT score and the distribution of the DT categories is different for each discipline. For example, MKA covers an even spectrum of DT categories whereas, not surprisingly, for TRH the vast majority of patients is associated to DT 4 since wounds and injuries often need quick care. Correspondingly, TRH also has a high DT score

The primary goal of our work is to increase the amount of patients served within their DT, thus served within the target time set by their surgeons. This goal can be achieved in three ways.

Firstly, it can be achieved by increasing capacity on the supply side by opening new ORs and hiring the additionally required personnel. Increasing existing OR capacities requires additional financial and spatial resources which in our setting are not readily available.

Secondly, it can be achieved by allowing more flexibility and using an open scheduling strategy. Allowing more flexibility in the schedule has advantages [279] and allows to better deal with occasional peaks in demand of single disciplines. Open scheduling is for the University Hospital Leuven, as for many other hospitals, not an option as it is important for the hospital to maintain a periodic and repetitive schedule. This allows surgeons to block certain weekdays for surgery while keeping other fixed days free for consultation, scientific work and teaching.

Thirdly, as considered in this research, it can be achieved by improving patient scheduling practices. We tested policies for patient scheduling that do not in-

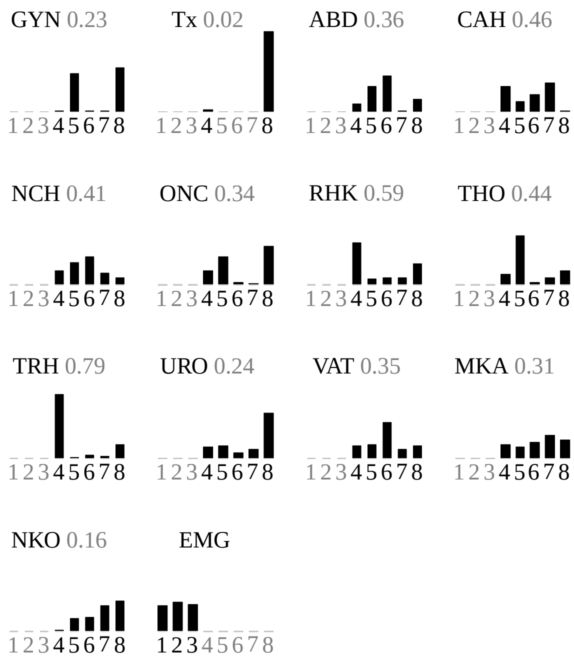


Fig. 1.2 The distribution of DT categories is markedly different for different disciplines. The number in gray denotes the DT score. 'EMG' stands for non-elective. The figure is based on data covering the entire years 2012-2013.

volve a computer. This is done as surgeons (or secretaries) at the University Hospital Leuven, and in Flanders (Belgium) in general [51], typically create patient schedules manually. Moreover, surgeons schedule their patients individually and therefore generally will not coordinate their schedules amongst each other.

We focus on the surgery to date assignment step. We do not sequence and do not determine the start time of surgeries. Those two factors are not important in our setting as elective patients are available the whole day and surgeons usually "own" an OR for the entire day. A surgery, therefore, starts directly once the preceding surgery is finished.

Additionally, we heavily focus on aspects that are important to get realistic results. This is true with regards to both the developed model and the tested scheduling methods. The model is realistic as we included all the aspects that we found to have a major effect on the results. This includes modeling aspects that relate to patient attributes (e.g., arrival, duration), to the structure of the setting (e.g., block assignment schema) and the processes (e.g., rescheduling, non-elective allocation schema). Also the tested scheduling methods are realistic as they reflect considerations or processes that are important in reality.

Components of the model were created on the basis of hard data. For aspects that were not covered by the data, we relied on the insights of our contacts in the hospital. They consist of a mix of people from the hospital that together have all the necessary experience. This includes, among others, the head surgeon, the head nurse, the responsible of the bed allocations and people from capacity management and the data gathering group.

Our contribution to the existing literature is therefore twofold. Firstly, we add to the sparsely addressed literature of dynamic advance surgery scheduling with stochastic arrivals. Secondly, we created a comprehensive model of the real hospital setting. We therefore modeled aspects of the real setting that are often not considered in the literature, but are important to include.

Chapter 2

Literature review

In hospitals, the operating room (OR) is a particularly expensive facility and thus efficient scheduling is imperative. This can be greatly supported by using advanced methods that are discussed in the academic literature. In order to help researchers and practitioners to select new relevant articles, we classify the recent OR planning and scheduling literature into tables regarding patient type, used performance measures, decisions made, OR up- and downstream facilities, uncertainty, research methodology and testing phase. Based on these classifications, we identify trends and promising topics.

2.1 Introduction

Within the hospital, considerable attention is given to ORs as they represent a significant segment of hospital costs [132]. Out of the many aspects of OR management, we focus our attention on planning and scheduling problems (we use the terms planning and scheduling interchangeably).

Given the importance of OR scheduling, it is not surprising that many research groups from the operations research community provide solution approaches to the problems that affect it. Reviews on this literature are important as they help researchers to select relevant articles for their research setting and serve as a guide for practitioners (e.g., hospital manager) to quickly find papers that can contain useful managerial insights. Additionally, reviews preferably help to identify promising practices and show recent trends (i.e., hot topics).

In order to cover these aspects, we define the following two research tasks. First, to classify the recent OR planning and scheduling literature (Sec. 2.3.1-2.3.7) using a simple, but comprehensive framework. For this task, we build upon the work carried out by Cardoen et al. [50] and Demeulemeester et al. [65]. Second, we look for evolutions over time, common approaches and relations between the different classification fields (Sec. 2.3.1-2.3.8).

The purpose of the remaining sections is to explain the research method (Sec. 2.2.1), to position this review in the existent group of reviews (Sec. 2.2.2) and to introduce the classification fields (introduction of Sec. 2.3).

2.2 Search method and other reviews

In Section 2.2.1, we introduce the procedure that we used to identify relevant articles. In Section 2.2.2, we discuss the structure and scope of reviews written on similar topics and position our review within the context of this existing literature.

2.2.1 Search method

We searched the databases Pubmed and Web of Science for relevant articles that are written in English and appeared between 2000 and 2014. Search phrases included combinations of the following words: surgery, case, operating, room, theatre(er), scheduling, planning and sequencing. We searched in both titles and abstracts and in addition checked the complete reference list of any already found article. As we endeavored to conduct the search process in an unbiased way, we believe we have obtained a set of articles that objectively represents the literature on OR planning. At the end of the search procedure, we identified 216

Table 2.1 The graphs showing trends are based on papers in the third column, while the tables additionally include the papers in the second column.

	2000-2003	2004-2014
Journal	24	137
Proceedings	3	42
Other	0	10
Total	27	189

technically oriented papers. Note that we chose to investigate trends only from 2004 onwards as in the preceding years not enough articles were published to get reliable results (Table 2.1).

We define an article as "technical" if it contains an algorithmic description of a method directly related to OR scheduling. Some articles are missing this algorithmic component and instead provide managerial insights. Those articles are excluded from the classification tables, as not all classification fields apply to them, but some of their insights are mentioned in the text. The quantitative descriptions provided in Sec. 2.3.1-2.3.8, which give insights into the changing trends set by the research community, are exclusively based on the technical contributions.

The majority of the included articles are recent publications (Fig. 2.1). This reflects the trend that the amount of published technical articles has been increasing significantly in the recent ten years.

We do not include topics related to business process reengineering, the impact of introducing new technologies, facility design or long-term OR expansion. Also, articles that deal with appointment scheduling are excluded from this review. This is the case as some of the basic assumptions that apply to appointment scheduling are not valid for surgery scheduling. For a review on appointment scheduling, we refer to Cayirli and Veral [54].

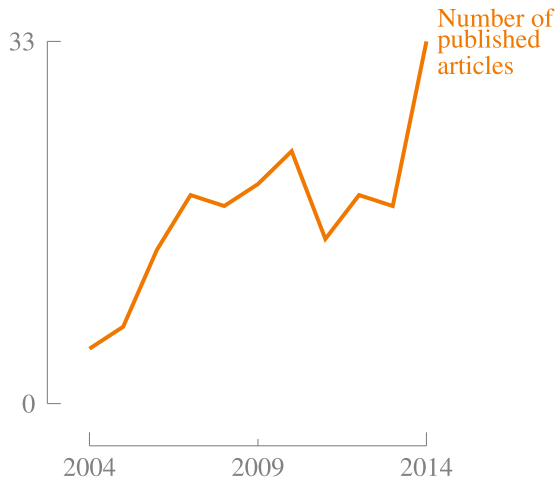


Fig. 2.1 The number of published technical articles in OR scheduling has been growing over the last decade.

2.2.2 Other reviews

In the past 60 years, a large body of literature on OR planning and scheduling has been published. The literature has been structured and reviewed by several authors, using a variety of classification techniques and frameworks. We grouped these reviews based on their scope and structure (Table 2.2).

Based on the scope of the literature reviews on ORs, we distinguish between three categories. The first category purely focuses on the OR department, including the post-anesthesia care unit (PACU). The second category targets the hospital in general, i.e., includes the intensive care unit (ICU), the ward [30] and patient flow planning in general and therefore discusses the OR as one of the areas that can be of interest in a hospital. The third category of reviews covers patient care services in general, such as ambulatory and surgical care [140].

In some of the literature reviews articles are classified based on the three hierarchical decision levels: strategic (long-term), tactical (medium-term) and operational (short-term). The strategic decision level involves decisions that affect both the number and the type of performed surgeries. The tactical level usually

Table 2.2 Existing reviews differ in their scope (*rows*) and classification structure (*columns*).

	Hierarchical categories	Custom categories
OR	[1, 116, 123]	[46, 65, 79, 91, 120, 181, 194, 226, 238]
Hospital	[30, 31, 36, 294]	[31, 36, 149, 254, 255, 277, 283]
Patient care services	[123, 139, 140]	[40, 106, 121, 123, 139, 140, 223]

Reviewing the literature according to hierarchal categories is a common approach. Articles appearing twice in the table use a multi-dimensional classification structure.

involves the construction of a cyclic schedule, which assigns time blocks to surgeons or surgeon groups. The final, operational level deals mostly with daily staffing and surgery scheduling decisions. Guerriero and Guido [116] also discuss papers that include a mix of the three levels. Similarly, Vissers et al. [294] propose a hierarchical framework for production control in healthcare. They distinguish between five levels and discuss for each level, amongst others, the type of decisions, the time horizon and the involved decision makers. With respect to the operational level, a further distinction can be made between off-line (i.e., before schedule execution) and on-line (i.e., during schedule execution) approaches [123].

In other literature reviews custom categories are used (Table 2.2). As such, Brailsford and Vissers [40] use the product life cycle stages to review 35 years of papers presented at the ORAHS conference. Moreover, Erdogan and Denton [91] review the literature according to the applied solution approach. Przasnyski [226] structures the literature based on general areas of concern, such as cost containment. Other reviews structure the literature on the basis of managerial or functional levels [223] and problem characteristics, e.g., the type of the arrival process [121].

Most literature reviews are not only reference points to articles, but also point out topics for future research. Guerriero and Guido [116] conclude that the three hierarchical levels are rarely studied together and argue that the tactical level has received increased attention in the last ten years. In contrast, Hans and Vanberkel [123] argue that future research should focus more on the tactical level.

Also, May et al. [194] make suggestions and argue that it might be promising to broaden the focus from operations research techniques to the economic and project management aspects of surgery scheduling. Additionally, Vissers et al. [294] suggest to put a larger emphasis on the multidisciplinary aspects of patient flow control systems and suggest to experiment with the effect of grouping patients in new ways, such as based on their length of stay (LOS) or surgery duration.

Furthermore, several authors emphasize that more research could be done on on-line rescheduling performed close to or on the day of surgery. Dexter et al. [79] provide a review on the few papers that include that type of decisions and emphasize the importance of the following four points: patient safety, open access to OR time, maximizing OR efficiency (defined as minimal overutilized OR time) and minimizing patient waiting time. Other reviews emphasize the need for more detailed models on the seasonality of demand, for more realistic constraints for surgeon and patient preferences and for a larger focus on the entire care pathway.

In this review, we propose a structure that is based on descriptive fields. We are not using hierarchical levels, since the boundaries between these levels can vary considerably for different settings and hence are often perceived as vague and interrelated [253]. Furthermore, this categorization seems to lack an adequate level of detail.

Moreover, other taxonomies that use one specific characteristic of the paper (e.g., solution technique) might prohibit the reader from easily finding a paper on a certain topic. For example, when a researcher is interested in finding papers on OR utilization, a taxonomy based on the solution technique does not seem very helpful. We think that the use of descriptive fields avoids these problems.

2.3 Descriptive fields

Each field analyzes articles from a different perspective, which can be either problem or technically oriented. In particular, we distinguish between seven fields:

- Patient characteristics (Sec. 2.3.1): reviewing the literature according to the elective (inpatient, outpatient) or non-elective (urgency, emergency) status of the patient;
- Performance measures (Sec. 2.3.2): discussing the performance measures (PM) such as utilization, idle time, waiting time, preferences, throughput, financial value, makespan and patient deferral;
- Decision delineation (Sec. 2.3.3): indicating what type of decision has to be made (date, time, room and capacity) and whether this decision applies to a medical discipline, a surgeon or a patient (type);
- Up- and downstream facilities (Sec. 2.3.4): discussing whether an approach includes other units (e.g., PACU and ICU);
- Uncertainty (Sec. 2.3.5): indicating to what extent researchers incorporate uncertainty (stochastic versus deterministic approaches);
- Operations research methodology (Sec. 2.3.6): providing information on the type of analysis that is performed and the solution or evaluation technique that is applied;
- Testing phase and application (Sec. 2.3.7): covering the information on the testing (data) of the research and its implementation in practice.

The structure we use is meant to balance between simplicity and comprehensiveness. It provides a simplified, but in our belief for the majority of the readers sufficiently accurate way to identify and select articles they are interested in.

The tables list and categorize all researched articles. Pooling them over the several fields enables the reader to reconstruct the content of a specific paper. They furthermore act as a reference tool to obtain the subset of papers that correspond to a certain characteristic.

Each section clarifies the terminology if needed and includes a brief discussion based on a selection of appropriate articles. At the end of each section, we discuss topics for future research.

Plots are provided for a selection of characteristics to point out the trends set by the research community. It should be noted that the percentages are calculated in relation to the total amount of technical papers. Also note that some fields are not interpretable for some methods and even though rare, some articles contain more than one single method. Moreover, the values for each year in the plots represent the average of the previous, the current and the next year. Using this moving average allows to spot larger research trends in an easier way. After all, a year with fewer publications does not imply that the topic has not been researched in that year.

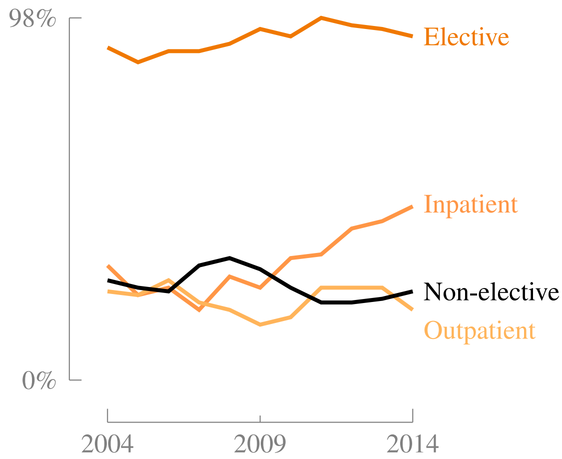
Finally, in the last part (Sec. 2.3.8) we go one step further and analyze the connection between different classification fields. This provides insights into research practices.

2.3.1 Patient characteristics

Two major patient classes are considered in the literature: elective patients and non-elective patients. The former class represents patients for whom the surgery can be planned in advance, whereas the latter class groups patients for whom a surgery is unexpected and hence needs to be fitted into the schedule on short notice. Although a consistent designation is lacking, a non-elective surgery is considered an emergency if it has to be performed immediately and an urgency if it can be postponed for a short time (i.e., days). As shown in Figure 2.2 and Table 2.3, the literature on elective patient scheduling is vast compared to its non-elective counterpart.

Although many researchers do not indicate what type of elective patients they are considering, some distinguish between inpatients and outpatients. Inpatients are

Fig. 2.2 The majority of articles relate to the elective patient. Contrary to what might be expected, the share of outpatient-related articles is not increasing. As some articles deal with both elective and non-elective patients, the sum of both values might add up to more than 100%.



hospitalized patients who have to stay overnight, whereas outpatients typically enter and leave the hospital on the same day.

In reality, there is an ongoing shift of services from inpatient to outpatient care (also called ambulatory care), which is reflected in a higher growth rate of the latter [8, 155, 195]. Moreover, according to the Milliman Medical Index, outpatient expenses increased on average by 9.9% over the years 2009-2013. This increase is largely attributed to increasing prices of existing and more expensive emerging services, but also to a relative increase in outpatient admissions [97, 199].

Compared to an inpatient setting, surgery in an outpatient setting has some particular features. For example, outpatient surgery often consists of more standardized procedures (e.g., routine surgeries, minimally invasive procedures). Moreover, since outpatients are not already present in a hospital ward before surgery, their actual arrival time is uncertain. These and other features might largely impact the choice of the scheduling technique. Despite the increasing importance of outpatient care in general, the share of articles on outpatient surgery remains flat (Fig. 2.2).

Besides planning electives, it is also important to consider non-electives. Non-

Table 2.3 The type of patient that is considered in articles is not always specified and, especially for the elective patient case, it is not always clear whether an inpatient or outpatient setting is researched.

Elective	
Inpatient	[2, 3, 14, 15, 16, 17, 24, 26, 28, 37, 39, 44, 45, 53, 55, 62, 64, 75, 94, 96, 103, 108, 111, 114, 122, 144, 147, 148, 157, 159, 169, 170, 178, 179, 180, 189, 190, 192, 197, 204, 205, 206, 217, 222, 227, 229, 231, 245, 256, 257, 262, 265, 268, 270, 276, 281, 282, 284, 285, 288, 295, 302, 303, 310, 311, 312, 313]
Outpatient	[15, 17, 27, 29, 39, 45, 48, 49, 67, 75, 76, 77, 83, 90, 96, 103, 107, 111, 113, 118, 119, 122, 137, 142, 148, 157, 159, 170, 173, 189, 190, 192, 204, 205, 206, 222, 230, 237, 242, 258, 262, 264, 265, 276, 282, 288, 296, 310]
Not specified	[4, 5, 9, 11, 12, 13, 18, 22, 38, 42, 43, 58, 60, 61, 63, 66, 68, 69, 72, 74, 80, 85, 92, 93, 95, 98, 99, 100, 101, 102, 104, 105, 110, 112, 120, 124, 125, 127, 128, 129, 131, 135, 138, 141, 143, 150, 151, 152, 156, 158, 161, 162, 163, 164, 165, 166, 168, 174, 182, 183, 184, 187, 188, 196, 200, 201, 202, 203, 209, 210, 211, 213, 214, 216, 218, 219, 220, 221, 225, 233, 235, 236, 241, 247, 249, 250, 259, 266, 267, 269, 271, 272, 278, 280, 290, 292, 300, 301, 304, 305, 306, 308, 309, 314]
Non-elective	
Urgent	[14, 38, 55, 95, 120, 122, 184, 202, 205, 208, 218, 222, 260, 314]
Emergent	[3, 14, 18, 37, 42, 45, 92, 93, 104, 105, 122, 128, 138, 147, 156, 161, 162, 163, 165, 166, 174, 188, 201, 204, 210, 218, 221, 222, 229, 260, 262, 264, 266, 267, 269, 282, 303, 304, 310]
Not specified	[158, 159, 213, 278, 301]
Unclear	[21, 23, 32, 33, 34, 56, 59, 70, 78, 117, 134, 145, 146, 175, 176, 185, 186, 246, 275, 286]

electives can be dealt within two ways. First, they can be incorporated in the elective schedule, which usually means that buffer capacity is reserved for them. For instance, van Essen et al. [92] explore the option of break-in-moments. A break-in-moment is the time point when an elective surgery is finished, presenting the opportunity to serve a waiting non-elective patient in the freed-up OR. In their setting, spreading these moments as evenly as possible over the day and ORs lowers non-elective waiting time. ORs are also shared between electives and non-electives in Lamiri et al. [165] who consider several stochastic optimization methods to plan elective surgeries. They present a solution method combining Monte Carlo sampling and mixed integer programming (MIP). They also test several heuristic methods from which the most efficient one proved to be tabu search.

Second, non-electives can be channeled into dedicated non-elective ORs. This requires however that a constant number of ORs is reserved for them and therefore leaves less free capacity for elective patients. Wullink et al. [304] show that this policy increases the waiting time for non-electives, while Heng and Wright [130] show that this decreases the number of elective cancellations and the amount of OR overtime. Recently, the combined effect of the use of dedicated ORs and a new policy for the urgency classification system is studied by a before-and-after study [171, 240].

A scenario where a hospital dedicates all of its ORs to emergency services is the case of a disaster. As a consequence, all elective surgeries are canceled while resources are redirected to provide quick care to non-electives. This type of non-elective patient is an urgency, as quick but not necessarily immediate care is required. Nouaouri et al. [208] sequence a large number of patients resulting from a disaster, with the objective of maximizing patient throughput. Their approach identifies patients that cannot be served by the given hospital and therefore have to be transported to another one.

Recently, Ferrand et al. [105] have researched a setting with a combination of dedicated and flexible ORs and show that it outperforms, in terms of patient waiting time and OR overtime, both the settings with shared ORs as well as the ones with dedicated ORs. The trade-off between patient waiting time and OR overtime represents the balance between an adequate degree of responsiveness to non-electives and the efficient use of OR resources.

Some authors use more than two urgency classes, i.e., they generalize the two category case of electives and non-electives. The highest urgency category may then be assigned to patients who need immediate care, whereas lower urgency categories can be assigned to patients who can wait for surgery for an extended period of time (e.g., months). For scheduling or evaluation purposes, each urgency category may be assigned a priority score [269] or a surgery target time [288].

An alternative way to categorize surgeries is on the basis of their discipline (e.g., cardiology) and surgery type (e.g., knee surgery or based on the ICD code). Surgery scheduling of different disciplines can to some extent be done independently, as the disciplines are often assigned to separate ORs. This is not the case for surgery types as one OR will typically accommodate more than one type of surgery. However, as a surgery type consists of surgeries that have a similar surgery duration, LOS and resource requirement (e.g., medical equipment), types are often used in models to formulate optimization problems in more general terms than what would be possible at the individual patient level.

For future research, more studies on outpatient surgery are needed. There is already a substantial amount of research on appointment scheduling in outpatient centers, but in most of this research schedules are created on the basis of appointment slots, neglecting the fact that surgery durations are highly variable.

Moreover, a mechanism that is not researched enough in the literature is patient bulking (i.e., patients deciding at any point in time to leave the waiting list). This can happen due to a variety of reasons (e.g., the patient decides to get surgery at another hospital). Patient bulking is important to model, but getting reliable real data on this mechanism can be challenging (e.g., hospitals might simply have no or only partial data available on patients that did not get surgery at their facility).

2.3.2 Performance measures

Different PMs emphasize different priorities and will favor the interests of some stakeholders over others. A hospital administrator could be interested in achieving high utilization levels and low costs, while medical staff might care less about cost factors and rather aim to achieve low overtime. The patient, as the client of the hospital, might care little about the above factors and only desires short waiting times.

Many authors in the scientific community try to find a compromise between the interests of different stakeholders and therefore simultaneously include several PMs. The most common approach is to include a weighted sum of these measures. We distinguish between the following major PMs: waiting time, utilization, leveling, idle time, throughput, preferences, financial measures, makespan and patient deferral.

As shown in Figure 2.3, patient waiting time is a frequently used PM. In this review, waiting time includes both direct waiting time (i.e., waits on the day of surgery) and indirect waiting time/access time (related to the size of the waiting list). Which one is used in an article can often be derived from the type of decision made in the model (Sec. 2.3.3). Wachtel and Dexter [298, 299] investigate the increase in waiting time on the day of surgery, for both surgeon and patient, caused by tardiness from scheduled start times. They conclude that the total duration of preceding cases is an important predictor of tardiness, i.e., the tardiness per case grew larger as the day progressed. A reduction of tardiness can be achieved by modifying the OR schedule to incorporate corrections for both the lateness of first cases of the day and the case duration bias.

Although surgeons are considered to be a valuable resource, their waiting time is included in a surprisingly low number of papers (Table 2.4). Part of the explanation is related to the fact that waiting time for the surgeon is mostly important in settings that are less frequently discussed in the literature (e.g., a setting where surgeons are allowed to book in any available slot).

We relate underutilization to undertime and overutilization to overtime, although they do not necessarily represent the same concept. Utilization refers to the workload of a resource, whereas undertime or overtime includes some timing aspect. Hence, it is possible to have an underutilized OR, which runs into overtime. In some articles it is unclear which view is applied. Therefore, we group underutilization with undertime and similarly overutilization with overtime.

Minimizing overtime is a popular objective (Fig. 2.3). This is not surprising as overtime in the OR can result in the dissatisfaction of the surgical staff, in high costs for the hospital (as higher wages typically apply for the time beyond the normal working hours), in surgery cancellations and in the disruption of the schedule in downstream departments. Dexter and Mario [81] establish that a correction of systematically underestimated lengths of case durations would not markedly reduce OR overutilization. They came to this conclusion as in their study few surgeries had a high probability of taking longer than scheduled. Tancrez et al. [266] propose an analytical approach that takes into account both stochastic surgery times and random arrivals of emergency patients. They show how the probability of running into overtime changes as a function of the total number of scheduled surgeries per day. Adan et al. [2] formulate an optimization problem that minimizes the deviation from a targeted utilization level for the OR, the ICU, the medium care unit and the nursing staff. The deviation is measured as the sum of overutilization and underutilization.

For some hospitals, measuring regular OR utilization is important. Interestingly, its use decreased from 2004 on until 2008, but stabilized from then on (Fig. 2.3). An example where the utilization of the surgical suit is maximized using an integer programming model and an improvement heuristic is provided by Marques et al. [189]. They schedule patients from the waiting list for the next week and assume that overtime is not allowed in the elective schedule. Luangkesorn et al. [177] argue against the use of utilization as a PM and argue that instead congestion metrics such as blocking and diversion should be used.

Figure 2.3 also shows that patient throughput is relatively rarely used. It is a quantitative measure, that is usually associated with the amount of patients that is served.

In contrast, preference-related measures most often cover some qualitative aspect. They experienced a peak of interest around 2010. Noteworthy is that both in general health care [133] and in the operations research literature value- and quality-based approaches seem to be getting increasingly important. For example, the preferences of cataract surgery patients of one surgeon are investigated by Dexter et al. [82]. The surgeon's patients place a high value on receiving care on the day chosen by them, at a single site, during a single visit and in the morning.

Preferences can also be embodied in patient priorities. Testi et al. [270, 272] define a model where the position of a patient on a waiting list is defined by a priority scoring algorithm, which considers both patient urgency (based on progression of disease, pain or dysfunction and disability) and time spent on the surgical waiting list. Clearly, priority scoring minimizes the total weighted waiting time of all patients. Therefore, an algorithm where patient priorities are equal, will minimize the average patient waiting time.

Including patient priorities drives OR scheduling in a more patient-oriented direction. Min and Yih [200] go one step further and explicitly incorporate an

Table 2.4 The division of articles based on the used performance criteria.

Waiting time	
Patient	[3, 9, 18, 29, 45, 56, 61, 64, 66, 67, 68, 95, 96, 104, 105, 107, 117, 118, 119, 120, 122, 134, 138, 141, 142, 145, 146, 152, 157, 166, 168, 169, 176, 201, 202, 203, 205, 209, 217, 219, 220, 221, 229, 231, 242, 245, 247, 249, 257, 258, 262, 266, 267, 268, 269, 270, 276, 282, 296, 304, 310]
Surgeon	[22, 58, 66, 68, 168, 182, 227, 288, 296, 311, 312, 313]
Leveling	
OR	[27, 44, 92, 186, 187, 209]
Ward	[24, 26, 28, 44, 53, 93, 94, 108, 124, 178, 179, 211, 241, 265, 284, 285]
PACU	[27, 48, 49, 93, 137, 184, 185, 246, 262, 280]
Patient volume	[183, 209, 265, 269]
Overutilization	
OR	[2, 3, 22, 29, 38, 42, 44, 45, 55, 56, 58, 59, 64, 66, 67, 68, 69, 70, 72, 78, 85, 90, 93, 95, 96, 98, 99, 100, 101, 102, 104, 105, 117, 118, 119, 120, 125, 135, 138, 142, 143, 144, 145, 146, 147, 151, 157, 161, 162, 163, 165, 166, 168, 170, 175, 182, 183, 184, 186, 188, 196, 197, 200, 201, 202, 209, 210, 214, 220, 221, 222, 225, 227, 229, 231, 233, 235, 236, 247, 249, 256, 265, 266, 267, 269, 271, 276, 278, 281, 295, 300, 303, 304, 308, 311]
Ward	[44, 55, 95, 295]
ICU	[2, 3, 64, 147, 214, 295]
PACU	[2, 3, 48, 49, 64, 90, 196]
Underutilization	
OR	[2, 3, 33, 34, 55, 58, 59, 64, 72, 98, 99, 100, 101, 102, 124, 135, 145, 146, 147, 151, 157, 163, 168, 170, 175, 188, 197, 209, 210, 214, 233, 250, 265, 269, 276, 280, 295, 300, 308, 310, 312, 313, 314]
Ward	[295]
ICU	[2, 3, 64, 147, 295]
PACU	[2, 3, 64, 267]
OR idle time	[29, 58, 66, 68, 96, 103, 111, 120, 131, 144, 169, 182, 188, 225, 227, 245, 311, 312, 313]
OR utilization	[9, 15, 17, 18, 23, 37, 38, 39, 45, 56, 60, 72, 75, 95, 104, 105, 107, 113, 125, 128, 148, 157, 166, 168, 179, 189, 190, 192, 209, 221, 247, 258, 262, 269, 271, 276, 278, 288, 304]
Throughput	[9, 15, 16, 17, 18, 23, 37, 44, 113, 128, 129, 148, 157, 170, 188, 190, 192, 197, 206, 208, 230, 241, 247, 258, 269, 271, 282]
Preferences	[4, 5, 16, 28, 32, 42, 48, 49, 60, 63, 83, 93, 114, 147, 158, 165, 178, 200, 201, 203, 213, 214, 217, 218, 231, 242, 259, 260, 265, 268, 270, 271, 272, 288, 290, 292, 301, 309]
Financial	[22, 32, 43, 59, 62, 69, 74, 75, 76, 77, 80, 85, 110, 120, 138, 159, 173, 176, 179, 180, 188, 204, 259, 286, 303]
Makespan	[11, 12, 13, 63, 98, 101, 102, 103, 111, 137, 164, 169, 170, 175, 184, 196, 222, 237, 242, 256, 275, 302, 305, 306]
Deferral/postponement	[3, 14, 38, 45, 59, 62, 64, 90, 93, 95, 112, 128, 131, 152, 156, 157, 174, 219, 220, 221, 229, 247, 262, 264, 271, 314]
Other	[2, 3, 16, 18, 21, 23, 56, 63, 64, 90, 93, 107, 108, 119, 124, 131, 141, 143, 158, 161, 162, 163, 173, 176, 179, 184, 187, 188, 196, 197, 202, 211, 216, 220, 222, 235, 236, 245, 257, 266, 267, 269, 290]

additional factor, namely the cost of OR overtime. In their model, if many high priority patients are on the waiting list, ORs will be kept open longer. This means that the surgery postponement costs are balanced against OR overtime costs. The authors establish that patient prioritization is only useful if the difference between the cost coefficients associated with different priority classes is high, as otherwise a similar schedule can be obtained by using the average postponement cost. Additionally, the relative cost ratio between the cost of patient postponement and OR overtime should not be low, as a low ratio would imply high overtime costs and therefore prioritizing would only marginally affect the surgery schedule.

An alternative and increasingly popular perspective on patient prioritization is the use of surgery target/DTs (e.g., knee surgeries need to be performed within two weeks). This is similar to their use in machine scheduling problems (e.g., [115]). DTs can be medically indicated, which entails that certain conditions will get worse if not dealt with in time. They therefore split the patients into various patient priority groups. As the importance of the waiting time of patients in different groups varies largely, a weighted formula can be used. The weight assigned to each priority group will need to reflect the urgency assigned to that group [239]. DTs can be set up by the authority of a larger geographic region such as a government [7, 20] or defined by a lower level authority such as a hospital [288].

Next to patient preferences or priorities, surgeon's preferences can be accounted for. As such, Meskens et al. [196] define the affinity between the staff members of the surgical team (i.e., surgeons, nurses and anesthesiologist). By including this measure into a multi-objective optimization procedure, they try to ensure that team members are working together with their preferred colleagues.

Some authors use purely financial objectives. In Stanciu and Vargas [259], protection levels (i.e., the amount of OR time reserved in a partitioned fashion for each patient class) are used to determine which patients to accept and which to postpone during the planning period under study. A patient class is a combination of the patient reimbursement level and the type of surgery. A patient class enjoys higher priority if its expected revenue per unit surgery time is higher. The goal of the method is to maximize expected revenues incurred by the sur-

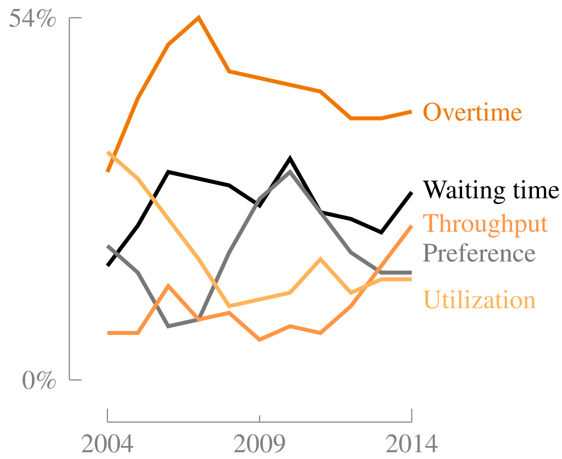


Fig. 2.3 Various performance measures are used in the literature from which the most popular is overtime. From 2008 onward, preference-related measures became increasingly popular, followed by a decline in interest after 2010.

gical unit. Patients, given their patient class, are accepted when the protection level for their class can accommodate them. The central question becomes how many requests to accept from low revenue patients and how much capacity to reserve for future high revenue patients.

Financial considerations are also expressed by Wachtel and Dexter [297], who argue that if OR capacity is expanded, it should be assigned to those subspecialties that have the greatest contribution margin per OR hour (i.e., revenue minus variable cost), that have the potential for growth and that have minimal need for a scarce resource such as ICU beds. Furthermore, Wang et al. [303] trade off the cost of opening an OR against the overtime cost for overbooking an OR that is already open. They develop a stochastic model that incorporates uncertain surgery durations, emergency demand and the risk of surgery cancellation.

Lee and Yih [169] minimize the makespan (completion time) of ORs by reducing delays in the patient flow. This is done by determining appropriate surgery starting times. Makespan in general defines the time span between the entrance of the first patient and the finishing time of the last patient in the OR. Since minimizing the makespan often results in a dense schedule, deviations from the plan can result in complications that require adjustments to the schedule. An example is the arrival of a non-elective patient to the hospital.

In the case of a non-elective arrival, it might be necessary to cancel an elective patient, who will consequently be served on a later day. Occasionally, if a non-elective patient cannot be served in a timely manner at the hospital, the deferral of the patient to another hospital can be initiated. General reasons for patient deferrals in one specific hospital are discussed by Argo et al. [10].

The trade-off between unused OR time and the cancellation rate of elective surgeries is investigated by Zonderland et al. [314] using queuing theory. In their setting, electives are canceled because arriving semi-urgencies are fit into the schedule. They also provide a decision support tool that assists the scheduling process of both elective and semi-urgent cases. Herring and Herrmann [131] examine the single-day, single-OR scheduling problem and balance the costs between deferring waiting cases and blocking higher priority cases. They provide threshold-based heuristics for OR managers that allow them to gradually release unused OR time in the days leading up to the day of surgery.

Another way to avoid cancellations is to level the utilization of units up- and downstream of the OR. For example, an overutilized PACU can block the OR, therefore prohibiting patients who have already completed surgery from leaving it. A blocked OR will impact succeeding elective surgeries, as they are either delayed or canceled. This situation can be avoided if the OR schedule is constructed in a way that the utilization of the units connected to the OR are leveled. For instance, Ma and Demeulemeester [178] maximize the number of expected spare beds and investigate bed occupancy levels at wards. The added benefit of leveling the utilization of units connected to the OR is a more balanced workload for the medical staff.

For future work, it could be interesting to increasingly include behavioral factors into the models as PMs (e.g., the satisfaction of staff, the booking behavior of surgeons in relation to the size of their waiting list, surgical team efficiency in relation to working hours). In order to include these factors, a deep understanding of the functioning of OR teams is needed that probably cannot be gained from data analysis only. Developing this understanding requires hospital involvement, which might not always be feasible.

Moreover, case studies are missing on how modeling assumptions (e.g., excluding or including emergency patients) are influencing different PMs. Those case studies are important as they help researchers to decide on which components (aspects of the real setting) are necessary to include in their model (e.g., include rescheduling if overtime is used as a PM).

2.3.3 Decision delineation

In the literature, various other terms are used to identify typical OR-related scheduling problems. Magerlein and Martin [181] distinguish between advance and allocation scheduling. Advance scheduling is the process of fixing a surgery date for a patient, whereas allocation scheduling determines the OR and the starting time or the sequence of the procedures on the planned day of surgery. For reasons of clarity, we recommend to call the former patient-to-date assignment problems and the later patient-to-room-and-time scheduling problems.

Within patient-to-date scheduling problems (advance scheduling), another distinction can be made between dynamic and static scheduling. Dynamic surgery scheduling refers to a setting where a patient is given a surgery date at consultation time, whereas in static surgery scheduling the patient is put on a waiting list. Patients on the list are then scheduled at once, e.g., at the end of each week. Dynamic scheduling can be used in settings where waiting lists are rarely used and waiting times are relatively short.

These two problems are handled differently in the literature from a methodological perspective. For the static problem, the hospital can use an algorithm that provides a schedule, i.e., the algorithm substitutes the scheduler. For the dynamic case, the hospital is usually using policies which the scheduler (e.g., assistant of surgeon) should consider in daily practice.

Another common distinction is made between block and open scheduling. In block scheduling, slots or blocks (i.e., a combination of an OR and a day) are typically allocated to a discipline or to a surgeon group. In the subsequent step, surgeons are only allowed to book cases into the blocks assigned to them. The suitability of this approach in various hospital settings is discussed by van Oostrum et al. [212]. In open scheduling, surgeons are not restricted to a block schedule and can therefore plan surgeries into an arbitrary OR.

In Table 2.5, we provide a matrix that indicates what type of decisions are examined, such as the assignment of a date (e.g., on Friday, February 25), a time, a room or an amount of capacity. The articles are further categorized according to the decision level they address, i.e., to whom the particular decisions apply. We distinguish between the discipline level (e.g., pediatrics), the surgeon level and the patient level. Papers that are categorized in the column or row with label 'Other' examine a wide variety of aspects. Examples are capacity considerations with regard to beds [174, 246], OR to ward assignment (i.e., $\mathrm{OR_i}$ to $\mathrm{Ward_j}$) [269], patient-to-week assignments [314] and different timing aspects, such as the amount of recovery time spent within the OR [13].

Using Table 2.5, problems that target each decision level can easily be identified. The discipline level unites contributions in which decisions are taken for a medical specialty or a department as a whole. Vansteenkiste et al. [288] propose a model to reallocate OR capacity between and within disciplines in such a way that patients are treated within their DT.

At the surgeon level, decisions can involve individual surgeons and also surgeon groups (e.g., all surgeons who perform hip replacement). In Denton et al. [69], surgeries consecutively carried out by one individual surgeon define a surgery block. Surgery blocks are subsequently assigned to ORs. The problem is formulated as a stochastic optimization model that balances the cost of opening an OR with the cost of overtime.

As Table 2.5 shows, a large part of the literature aims at the patient level. At this level, the decision variables are formulated on the basis of the individual patient or the patient type (e.g., ICD-code).

In Fei et al. [102] patients are scheduled in two stages. In the first stage, patients are assigned to days and rooms, while in the second stage the exact daily sequence (timing aspect) is determined. This is a common way of scheduling patients, as the assignment of the day and the room for a given surgery is easier planned ahead in time than the exact starting time of the surgery, which is often only fixed close to the actual surgery date.

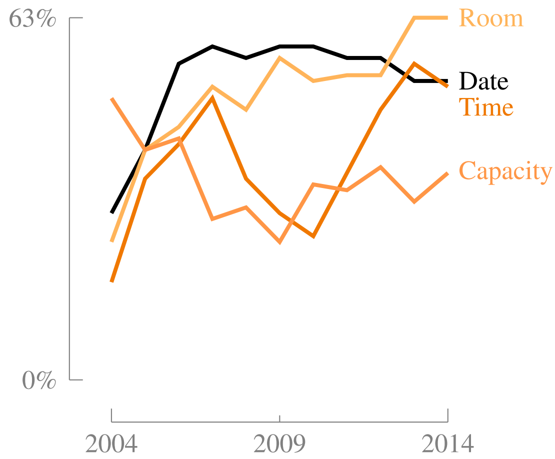


Fig. 2.4 Room assignment problems are increasingly popular in the literature. The interest in the time assignment step (e.g., sequencing) shows a more variable pattern, e.g., it has lost some of its popularity around 2010, but regained it towards 2014.

A problem setting where a date and a room (e.g., OR 1, OR of type B) are assigned to patients is discussed by Gomes et al. [113]. Their optimization method includes a component that predicts the duration of surgeries. This is important as the variance in surgery durations has a large impact on OR performance.

Time-related decisions can either relate to problems where a sequence (e.g., patient A follows B) or an exact surgery start time (e.g., 2.10 pm) is determined. A method to determine the latter is discussed by Schmid and Doerner [245] who show that it is beneficial to couple routing (e.g., transport from an examination room to the OR) and scheduling decisions.

Capacity-related decisions mainly focus on assigning OR time to disciplines, which often results in the Master Surgery Schedule (MSS), this is a 1 or 2 week cyclic plan where to each weekday and OR a specific discipline or surgeon is assigned. The construction of such an MSS is tested with three different policies by Cappanera et al. [44] who compare the efficiency (i.e., maximize throughput), the balancing effect (i.e., have a fair allocation of workload for all departments) and the robustness (i.e., prevent disruptions) of the resulting schedule. They also compare the performance of their policies in various hospital settings. Two models are presented by Manmino et al. [183] where, in the first model,

Table 2.5 The division of articles based on the decision (*columns*) and assignment (*rows*) level.

	Discipline level	Surgeon level	Patient level	Other
Date	[16, 17, 24, 33, 34, 43, 44, 56, 59, 61, 74, 108, 120, 134, 183, 241, 247, 270, 271, 284, 285, 310]	[14, 17, 26, 27, 28, 45, 53, 62, 103, 144, 151, 179, 219, 269, 281]	[2, 3, 4, 5, 14, 16, 42, 44, 45, 55, 56, 60, 64, 74, 75, 85, 95, 98, 99, 100, 101, 102, 110, 112, 113, 117, 118, 119, 120, 124, 125, 129, 138, 144, 145, 146, 147, 151, 152, 156, 157, 158, 161, 162, 163, 165, 170, 175, 178, 179, 180, 189, 190, 192, 197, 200, 201, 203, 209, 211, 214, 217, 219, 220, 222, 225, 229, 231, 233, 235, 236, 247, 249, 250, 257, 258, 265, 268, 270, 271, 272, 280, 281, 290, 292, 300, 308, 309]	[62, 85, 94, 95, 179, 268, 295]
Time	[16, 17, 24, 44, 56, 74, 120, 128, 134, 183, 247, 271]	[14, 17, 22, 26, 27, 28, 45, 62, 66, 144, 196, 281]	[4, 11, 12, 13, 14, 16, 22, 29, 44, 45, 48, 49, 56, 66, 67, 68, 70, 74, 85, 90, 93, 96, 98, 101, 102, 105, 111, 113, 118, 120, 127, 128, 129, 137, 141, 142, 143, 144, 145, 146, 158, 164, 166, 168, 169, 175, 182, 184, 185, 187, 189, 190, 192, 196, 197, 208, 222, 227, 231, 235, 236, 242, 245, 247, 256, 264, 271, 275, 281, 292, 302, 305, 306, 311, 312, 313]	[22, 23, 62, 85, 196, 245, 302]
Room	[16, 17, 33, 34, 44, 56, 59, 61, 108, 114, 134, 183, 241, 247, 270, 271, 284, 285, 310]	[17, 22, 27, 28, 53, 62, 69, 103, 144, 151, 179, 196, 219, 269, 281]	[4, 5, 16, 22, 39, 42, 44, 48, 49, 55, 56, 60, 63, 70, 72, 78, 85, 96, 98, 99, 100, 101, 102, 104, 105, 111, 113, 117, 119, 124, 125, 127, 129, 142, 143, 144, 145, 146, 151, 157, 158, 161, 163, 164, 170, 175, 179, 182, 184, 186, 187, 189, 190, 192, 196, 201, 203, 208, 209, 211, 214, 217, 218, 219, 220, 222, 227, 229, 231, 233, 235, 236, 237, 245, 247, 249, 250, 256, 258, 268, 270, 271, 280, 281, 292, 300, 302, 303, 304, 306, 308, 311, 312, 313]	[22, 62, 85, 179, 196, 245, 268, 302, 303]
Capacity	[16, 17, 37, 38, 43, 44, 56, 59, 61, 74, 120, 122, 128, 135, 148, 241, 247, 262, 271, 288, 310]	[17, 22, 32, 45, 58, 62, 66, 76, 77, 80, 151, 159, 179, 196, 219]	[2, 3, 5, 16, 22, 38, 44, 45, 56, 64, 66, 74, 95, 117, 120, 124, 128, 131, 138, 141, 147, 151, 173, 178, 179, 180, 196, 200, 202, 204, 213, 219, 220, 247, 259, 266, 267, 271, 286, 303, 314]	[22, 23, 62, 83, 95, 107, 174, 176, 179, 196, 205, 206, 216, 221, 230, 246, 282, 303]
Other	[278]	[219, 281]	[9, 13, 63, 92, 93, 95, 111, 119, 150, 169, 209, 219, 227, 245, 257, 260, 268, 281, 296, 301, 305, 314]	[83, 95, 245, 268, 276]

For example, articles dealing with the sequencing problem are found in column 3 and row 2 (header rows/columns are excluded). Articles dealing with advance scheduling are found in column 3 and row 1. Allocation scheduling models are generally found in column 3 and rows 2 and 3. Defining patient capacity requirements for a given day of the week are articles found in column 3 and both row 1 and row 4.

OR overtime is minimized and, in the second model, patient queue lengths are balanced amongst different specialties. For the second model they additionally develop a light robustness approach [109] that copes with the demand uncertainty.

Capacity problems can generally be solved in two ways. A hospital can either decide on the number of OR-days to assign to each specialty or, as is proposed by Testi et al. [271] and Adan et al. [3], it can decide on the number of patients it allocates to each OR session. Generally, the division of OR block time is a heavily constrained problem as different factors, such as the available OR block size (e.g., 9 hours), are taken into account. Performance measures that are used to drive such a model are among others the expected costs related to undertime and/or overtime and the number of unscheduled patients [59].

A capacity problem is also discussed by Masursky et al. [193] who forecasted long-term anesthesia and OR workload. They conclude that forecasting future workload should be based on historical and current workload-related data and advise against using statistical data on the local geographical population. The problem of forecasting workload is also addressed by Gupta et al. [122]. In their case study, simulation is used to answer capacity-related questions. They concluded that a one-time infusion of capacity in the hope to clear backlogs will fail to reduce waiting times permanently, while targeting extra capacity to highest urgency categories reduces waiting times for all categories, including those of low urgency patients. In situations where arrival rates increased, even if only within a specific urgency class, waiting times increased dramatically and failed to return to the baseline for a long time.

We think that there are two main advantages of identifying papers using the structure of Table 2.5 over an approach that is based on terminology. First, there will be problem settings that do not have a commonly used term and, second, different authors might use the same terminology for variants of the same problem. For instance, Fügener et al. [108] define an MSS as a discipline-to-date-and-room assignment, whereas in Banditori et al. [16] it is defined as a patient-to-date-room-and-capacity assignment. Table 2.5 provides therefore a less ambiguous way to identify certain problem settings.

We noted that there are many advanced and complex methods on static surgery scheduling. However, in some hospital settings patients have to be scheduled dynamically, which requires other methods [239].

For future work, it would be interesting to see more research on dynamic scheduling. Dynamic scheduling methods are already heavily used in the appointment scheduling literature. The reason they are scarcely used in the surgery scheduling literature is twofold. First, in many hospitals surgeries are scheduled statically, evidently requiring static methods. Second, the methods that are used for dynamic scheduling in an appointment setting are not easily transferable to a surgery scheduling setting for various modeling reasons, such as the fact that estimated slot durations in the former setting are assumed to be of equal length, while in the latter they are highly variable.

2.3.4 Up- and downstream facilities

As OR planning and scheduling decisions affect departments throughout the entire hospital, it seems useful to use an integrated approach and therefore incorporate upstream (e.g., outpatient clinic) and downstream facilities (e.g., the ICU or the PACU) in the OR scheduling process and as such to improve their combined performance. When this is ignored, we believe that improving the OR schedule may worsen the efficiency of those related facilities. Whether an article discusses an integrated or an isolated approach can be looked up in Table 2.6.

The ratio of articles that deal with the OR in an integrated way is staying around the 50% mark throughout the years 2004-2014 (Fig. 2.5). This is surprising as models are getting more complex and one would expect to observe an increasing interest in integrated approaches. One explanation for this lack of increase is the fact that we exclude articles that do not consider any type of OR planning. Therefore, articles that only deal with up- or downstream units, but do not take the OR explicitly into account, are not shown.

As shown in Figure 2.5, the problem of the congested PACU received more attention from 2007 onwards. If the PACU is congested, patients are not allowed

Table 2.6 In an integrated OR, upstream and/or downstream facilities such as the ICU, the PACU and the wards are considered.

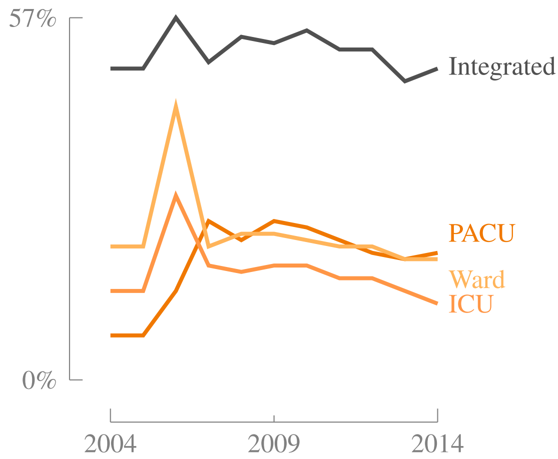
Isolated OR	[4, 5, 11, 21, 22, 29, 33, 34, 38, 42, 56, 58, 59, 60, 61, 62, 63, 66, 68, 69, 70, 72, 74, 75, 78, 80, 83, 85, 86, 92, 99, 100, 101, 103, 104, 105, 111, 112, 113, 117, 119, 122, 125, 129, 131, 134, 135, 138, 144, 150, 151, 152, 157, 158, 159, 161, 162, 163, 165, 166, 168, 170, 173, 174, 175, 176, 182, 183, 186, 187, 189, 190, 192, 200, 202, 203, 208, 209, 210, 213, 216, 217, 218, 219, 221, 225, 227, 229, 231, 233, 235, 236, 237, 245, 247, 249, 250, 258, 259, 260, 264, 266, 272, 275, 276, 278, 286, 288, 292, 296, 300, 301, 303, 304, 308, 309, 311, 312, 313, 314]
Integrated OR	[2, 3, 9, 12, 13, 14, 15, 16, 17, 18, 23, 24, 26, 27, 28, 32, 37, 39, 43, 44, 45, 48, 49, 53, 55, 64, 67, 76, 77, 90, 93, 94, 95, 96, 98, 102, 107, 108, 110, 114, 118, 120, 124, 127, 128, 137, 141, 142, 143, 145, 146, 147, 148, 156, 164, 169, 178, 179, 180, 184, 185, 188, 196, 197, 201, 204, 205, 206, 211, 214, 220, 222, 230, 241, 242, 246, 256, 257, 262, 265, 267, 268, 269, 270, 271, 280, 281, 282, 284, 285, 290, 295, 302, 305, 306, 310]

to enter it and are therefore forced to start their recovery in the OR itself, keeping it blocked. Iser et al. [143] use a simulation model to tackle this problem and compare OR overtime to PACU-specific PMs. Augusto et al. [13] show, using a mathematical model, the benefits of preplanning the exact amount of recovery time a patient will spend in the OR. Generally, as is typical for highly utilized systems, there is a sensitive relationship between overall case volume, capacity (of the PACU) and the effect on waiting time (to enter the PACU). This relationship is described in more detail by Schonmeyr et al. [246] using queuing theory.

The relationship between the ICU and the OR has been scarcely addressed in the last decade (Fig. 2.5). Kolker [156] reduces the number of patients served in another than their designated ICU to an acceptable level and defines the maximum number of elective surgeries per day that are allowed to be scheduled along with emergency arrivals. Litvak et al. [174] go a step further and tackle the ICU capacity problem in a cooperative framework. In their model, several hospitals of a region jointly reserve a small number of beds in order to accommodate emergencies and achieve an improved service level for all patients.

Similarly, also the bed management in the wards is closely related to the OR schedule and, in particular, to the MSS. In some hospitals, specialties need to ensure that they have enough capacity in their own wards in order to prevent

Fig. 2.5 An integrated OR planning and scheduling process is considered in around 50% of articles. The downstream facilities (i.e., PACU, ward, ICU) are the most common included units. As only the three main downstream facilities are shown, their count does not necessarily sum up to the total number of the integrated approaches.



bed misplacements, unnecessary movements between wards and OR blocking due to bed unavailability. Beliën and Demeulemeester [24] and Vanberkel et al. [284] for instance optimize the MSS in order to level the expected ward occupancy with a mathematical program (MP) and an analytical model respectively. More generally, Fügener et al. [108] propose an MSS that minimizes the cost of downstream units (i.e., capacity costs and staffing costs). The main idea in these three papers [24, 108, 284] is that based on the MSS, the expected workload in the wards can be calculated. This is the case as the probability distribution of arrival times in downstream units is known. This expected workload can bring possible resource conflicts to light, which then can be corrected by modifying the MSS.

Integrated approaches can also incorporate preoperative units. For example, Huschka et al. [142] consider both an intake and a recovery area as part of a simulation model of an outpatient procedure center. They test several daily scheduling and sequencing heuristics and investigate their impact on the average patient waiting time and the OR overtime. The authors found that these PMs are more influenced by the scheduled arrival time of patients and less by their sequence.

Recently, the integration of the OR schedule with alternative aspects gained attention, e.g., the combination of nurse rostering and OR scheduling [281, 306] and the inclusion of surgery scheduling into a broader perspective of the patient care process [110, 141].

For future work, we encourage the following topics on integration as they seem to have received only limited attention: the integration with upstream processes (e.g., preoperative assessment), the integration of the outpatient and inpatient surgery schedules (preferably also taking into account the outpatient clinic sessions), the integration of geographically dispersed OR units and the integration of other functional departments dealing with the ORs (e.g., the logistics department handling inventory, the financial department for reimbursements, the central sterilization services handling instrument sterilization). These topics are challenging, among others, because different departments/stakeholders are involved and therefore just collecting the necessary data can already be difficult.

2.3.5 Uncertainty

One of the major problems associated with the development of accurate OR planning and scheduling strategies is the uncertainty inherent to surgical services. Deterministic planning and scheduling approaches ignore uncertainty, whereas stochastic approaches explicitly incorporate it. In Table 2.7, we classify the articles according to the type of uncertainty that is incorporated.

As shown in Figure 2.6, stochasticity in the form of uncertain patient arrivals and surgery durations is frequently incorporated. Non-elective patient arrivals are in most cases impossible to predict in advance and additionally occupy a random amount of OR time, which often leaves OR managers with no option but to reserve capacity for them [266]. In contrast, the arrival of elective patients to ORs contains little uncertainty and is frequently considered as deterministic in the literature. If we narrow down the literature to contributions that explicitly incorporate non-elective patients, we see that around 80% of them use methods that incorporate some sort of uncertainty.

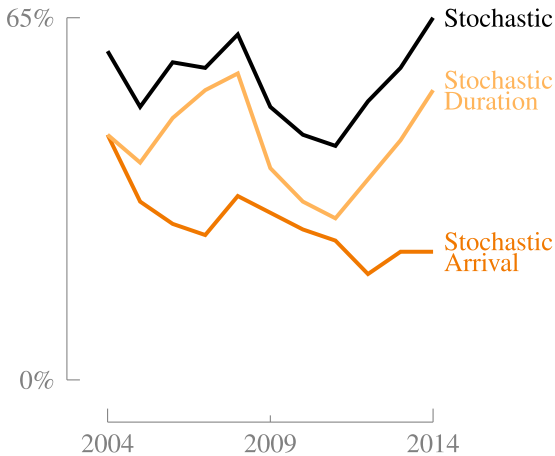
Table 2.7 In articles, stochasticity is frequently taken into account.

Deterministic	[4, 5, 12, 13, 14, 17, 21, 26, 27, 32, 33, 34, 43, 48, 49, 55, 56, 60, 62, 63, 76, 78, 83, 85, 86, 90, 93, 98, 99, 100, 101, 102, 104, 107, 110, 111, 113, 117, 127, 129, 135, 137, 141, 143, 144, 145, 146, 147, 151, 158, 159, 164, 173, 175, 180, 183, 187, 189, 190, 192, 196, 203, 204, 208, 209, 213, 214, 217, 219, 222, 227, 231, 233, 235, 236, 237, 241, 245, 256, 257, 258, 265, 268, 270, 272, 275, 280, 281, 284, 285, 288, 290, 292, 295, 302, 305, 306, 308, 309]
Stochastic	
Arrival	[3, 9, 18, 24, 28, 38, 39, 45, 59, 61, 64, 74, 75, 92, 95, 103, 105, 112, 114, 120, 122, 128, 131, 152, 156, 157, 161, 162, 163, 166, 176, 178, 179, 188, 200, 202, 205, 206, 216, 218, 220, 221, 229, 230, 247, 249, 259, 260, 262, 266, 267, 269, 271, 282, 286, 301, 303, 304, 310, 314]
Duration	[2, 3, 9, 11, 15, 16, 18, 22, 23, 24, 28, 29, 38, 42, 44, 45, 58, 59, 64, 66, 67, 68, 69, 70, 72, 74, 92, 95, 96, 105, 112, 114, 118, 119, 120, 122, 124, 125, 128, 134, 138, 142, 152, 157, 161, 162, 165, 166, 168, 169, 170, 176, 178, 182, 184, 185, 186, 188, 197, 200, 201, 202, 205, 206, 210, 211, 216, 218, 221, 225, 229, 230, 242, 246, 247, 249, 250, 262, 264, 266, 267, 269, 271, 276, 278, 282, 286, 296, 300, 301, 303, 304, 310, 311, 312, 313]
Other	[9, 11, 16, 18, 42, 44, 45, 77, 80, 108, 138, 148, 150, 170, 174, 178, 179, 182, 201, 230, 267, 301]

Surgery durations are difficult to predict because for some surgeries the magnitude of the procedure only becomes apparent once the surgery is already in progress. Additionally, the durations often depend on various complex factors, e.g., the characteristics of the patient, the surgeon and the surgical team. As individual surgery durations are uncertain, also their sum, the total workload per OR, is uncertain. Out of all papers, 44% takes duration uncertainty into account, while 28% consider arrival uncertainty.

Duration uncertainty is a central element in Denton et al. [69] as well as in Batun et al. [22]. In Denton et al. [69], decisions include the number of ORs to open and assignments of surgery blocks to ORs, whereas in Batun et al. [22] also the sequence of patients and the starting time of surgeons is determined. Both models aim at minimizing OR opening and OR overtime costs, where Batun et al. [22] additionally consider surgeon idle times. The functional difference between their methods lies in the way surgery to OR assignments are carried out. In Denton et al. [69], the common practice of assigning a surgery block to a single surgeon (block scheduling) is followed, whereas Batun et al. [22] consider the scenario of pooled ORs where surgeons are allowed to switch between ORs.

Fig. 2.6 Some type of uncertainty is taken into account in more than half of the papers.



OR pooling allows to carry out surgeries in parallel as the main surgeon only needs to be present during the critical part of the surgery and can move to the next patient before closing the patient.

Shylo et al. [250] introduce a chance-constrained model of overtime that, based on the normal approximation of the sum of durations in one OR-day, provides near-optimal solutions to the surgery to time block assignment problem. Using real data, they show that the developed algorithm is particularly suitable for specialties with high patient volumes per OR-day.

Surgery rescheduling limits the impact that deviations from the initial OR schedule have on the hospital. These deviations on the day of surgery are caused by an uncertain workload due to possible emergency arrivals, deviations from the estimated surgery durations or variable LOS in downstream units. Other causes that can lead to deviations include staff unavailability, equipment failure, late arrival of patients or staff and, in an outpatient setting, patient no-shows. To limit the impact, interventions throughout the day in the form of rescheduling might be needed.

We distinguish between two main types of interventions: cancellations and OR reassignments. In case of an OR reassignment, the patient is still served on the planned day, but is moved or rescheduled to another OR. A more severe intervention is when a patient cannot be served on the planned day and needs to be canceled. This patient will need to be fitted into the elective schedule of another day. Cancellations are performed throughout the day [88, 239] and can vary considerably from setting to setting (e.g., Leslie et al. [172] (8%), Xue et al. [307] (18%), Epstein and Dexter [88] (11.8%) and Samudra et al. [239] (3.4%)). Many papers report scheduling issues as one of the main causes for case cancellations, next to medical reasons and preoperative or structural reasons [57, 172, 307]. This emphasizes the need for good proactive and reactive scheduling approaches.

An optimization model is proposed by Stuart and Kozan [264] for rescheduling patients on the day of surgery. Their model resequences elective and non-elective patients in each OR whenever a surgery is completed. Using a branch-and-bound algorithm, they maximize the weighted throughput. This implicitly minimizes the patient cancellation rate. Similarly, Erdem et al. [90] reschedule elective patients upon the arrival of an emergency patient. Considering both the OR and the PACU, they minimize the cost of disruptions using a MIP and a genetic algorithm. A decision support system is provided by van Essen et al. [93], where in reaction to disruptions in the schedule adjustments are proposed to the OR manager. An MP is used to derive the decision rules, e.g., either shifting a surgery or scheduling a break between two surgeries.

A method where surgeries are rescheduled across multiple ORs is introduced by Zheng et al. [311]. In their method, at each time point when an OR becomes unoccupied it is determined which surgery to start next. This decision is based on the surgeon's waiting time as well as the OR's idle time and overtime.

It should be clear that operations research techniques are able to deal with stochasticity and especially simulation techniques (used in around 50% of the stochastic literature) and analytical procedures (used in around 20% of the stochastic literature) seem to be well suited. Stochastic programming (e.g., two-stage linear programming) can also be useful to solve these problems. However, there are a limited number of papers that use this technique to solve real-life problems. This constitutes an area for future research.

Studies mostly assume a certain level of variability, based on analyzing historical data, and use this information as input for models. Unfortunately, only limited attention is paid to the reduction of variability within the individual processes. As an example, consider the estimation of surgery durations. Instead of immediately determining the distribution of the surgery durations, one could examine first whether the population of patients for which the durations are taken into account is truly homogeneous. If not, separating the patient population may result in a decreased variability even before the planning and scheduling phase is executed. Since the estimation of surgery durations exceeds the scope of this literature review, we do not elaborate further on this issue. Another example is the reduction of turnover times, as discussed by Kodali et al. [154].

For future work, we encourage further research on policies that reschedule patients on the day of surgery itself or even on the days before the surgery. Rescheduling is an important mechanism that affects both patient and staff satisfaction.

2.3.6 Operations research methodology

The literature on OR planning and scheduling exhibits a wide range of methodologies that fit within the domain of operations research. Table 2.8 provides an overview of the techniques that are used to solve OR planning and scheduling problems.

In some approaches the impact of specific changes to the problem setting is examined. We refer to such an approach as a scenario analysis since multiple scenarios, settings or options are compared to each other with respect to the PMs. Performing a scenario analysis is popular (Fig. 2.7) and especially in the DES modeling literature often done.

An integrated DES model is introduced by Steins et al. [262], in which preoperative care and a PACU are considered. The arrival of case types, the surgery time and the LOS in the PACU are represented as probabilistic distributions.

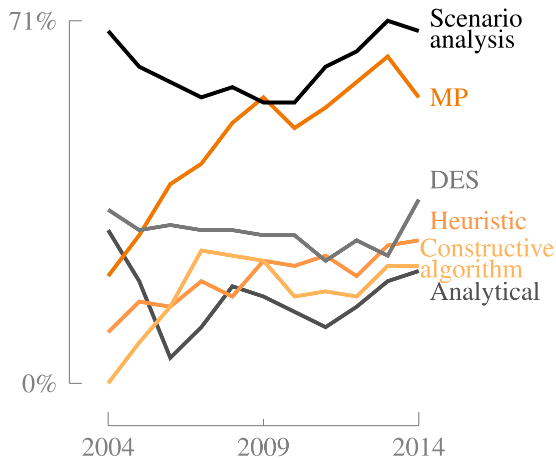
An analytical approach, using a Markov model, is introduced by Tancrez et al. [266] who determine the amount of OR capacity needed to accommodate non-

Table 2.8 There are different solution techniques used in the literature.

Simulation	
Discrete-event	[3, 9, 11, 14, 15, 16, 18, 23, 29, 37, 38, 39, 44, 45, 64, 72, 74, 75, 92, 95, 96, 102, 103, 104, 105, 107, 118, 122, 128, 142, 143, 148, 152, 156, 157, 166, 170, 174, 176, 178, 179, 184, 185, 186, 188, 197, 201, 205, 206, 218, 220, 221, 227, 230, 242, 247, 249, 250, 262, 269, 271, 276, 282, 304, 310]
Monte-Carlo	[38, 67, 77, 125, 161, 162, 165, 168, 201, 202, 209, 216, 313]
Mathematical programming	
Linear programming	[13, 66, 76, 77, 159, 204, 225, 311]
Goal programming	[2, 3, 32, 64, 214, 265]
Integer programming	[4, 33, 34, 42, 48, 53, 56, 59, 62, 83, 93, 94, 113, 131, 144, 179, 180, 208, 211, 233, 241, 242, 258, 268, 269, 270, 271, 286, 312]
Mixed integer programming	[16, 17, 22, 24, 28, 29, 44, 69, 85, 92, 110, 119, 124, 134, 141, 145, 146, 147, 150, 151, 161, 162, 165, 179, 182, 183, 189, 201, 203, 217, 219, 220, 222, 227, 229, 236, 281, 292, 302, 310]
Column generation	[98, 100, 101, 102, 111, 124, 129, 134, 161, 163, 164, 211, 280, 303]
Branch-and-price	[26, 49, 99, 179, 180]
Dynamic programming	[12, 13, 26, 49, 98, 99, 131, 138, 163, 200, 202, 286, 308]
Other	[12, 13, 24, 28, 44, 61, 80, 127, 187, 196, 225]
Improvement heuristic	
Simulated annealing	[24, 28, 55, 67, 92, 94, 108, 125, 165, 166]
Tabu search	[63, 92, 98, 137, 165, 206, 242]
Genetic algorithm	[60, 90, 92, 96, 98, 102, 118, 169, 192, 235, 236, 256, 257, 275, 300, 309]
Other	[33, 34, 59, 61, 68, 92, 96, 108, 125, 161, 163, 165, 166, 186, 189, 190, 201, 231, 233, 245, 250, 259, 305, 306, 308, 312]
Constructive algorithm	[5, 11, 12, 24, 28, 29, 42, 56, 66, 68, 69, 72, 74, 78, 92, 100, 101, 102, 117, 125, 131, 135, 138, 142, 143, 158, 161, 163, 164, 165, 166, 169, 175, 190, 196, 202, 227, 230, 233, 249, 268, 269, 288, 292, 303, 312]
Analytical procedure	[29, 38, 58, 59, 61, 62, 66, 86, 101, 108, 112, 114, 120, 138, 162, 165, 170, 174, 176, 200, 202, 210, 213, 218, 246, 260, 266, 267, 278, 286, 296, 301, 314]
Branch-and-bound	[48, 69, 108, 209, 264, 295]
Scenario analysis	[3, 4, 5, 9, 11, 13, 14, 15, 16, 17, 18, 23, 27, 29, 32, 37, 38, 39, 42, 44, 45, 59, 60, 62, 63, 64, 66, 67, 68, 69, 70, 72, 74, 75, 76, 77, 78, 85, 90, 93, 94, 95, 96, 100, 101, 102, 103, 105, 107, 112, 113, 114, 118, 119, 122, 125, 128, 131, 134, 135, 137, 142, 146, 148, 152, 156, 157, 159, 164, 166, 168, 169, 170, 173, 174, 176, 178, 179, 182, 183, 184, 185, 186, 188, 189, 196, 197, 201, 202, 203, 204, 205, 206, 209, 210, 214, 216, 217, 218, 220, 221, 227, 229, 230, 233, 236, 241, 242, 246, 247, 249, 250, 257, 258, 259, 262, 266, 267, 269, 271, 272, 276, 278, 280, 281, 282, 284, 285, 286, 288, 295, 296, 301, 302, 303, 304, 310, 311, 312]

There are a few papers that are not mentioned in the table as they include a method that could not be clearly assigned to any of these categories.

Fig. 2.7 From the major solution techniques used in the literature only MP experienced a strong growth in popularity.



elective patients. Simulation is used to show that the assumptions required to build the Markov chain have a minor influence on their final analytical results. In their work, the stochasticity in OR capacity is the consequence of randomly arriving non-elective patients occupying an uncertain amount of OR time.

Even without non-elective patient arrivals, it might be difficult to predict the required OR capacity on a day, as surgery durations are unknown in advance and can vary considerably in length. Olivares et al. [210] analytically investigate the decision-making process of reserving OR capacity using the newsvendor model. In their approach, an estimate is given of the cost placed by the hospital on having idle capacity and the cost of a schedule overrun. Their results reveal that the hospital under study places more emphasis on the costs of having idle capacity than on the costs of a schedule overrun and long working hours for the staff.

Table 2.8 shows that MPs, improvement heuristics and constructive algorithms are frequently used. As opposed to DES and analytical models, MPs, such as MIPs, deal with combinatorial optimization problems.

In a large number of cases, the objective function of the optimization model includes under/overtime or under/overutilization. Those PMs are rarely used by themselves, but are usually part of a multi-objective formulation as two thirds of MP models use multiple objectives.

In most of the MPs, the decision applies to the elective patient, as in Min and Yih [201]. In their work, a stochastic MIP model is proposed and solved by a sampling-based approach. The surgery durations, the LOS, the availability of a downstream facility and new demand are assumed to be random with known distributions.

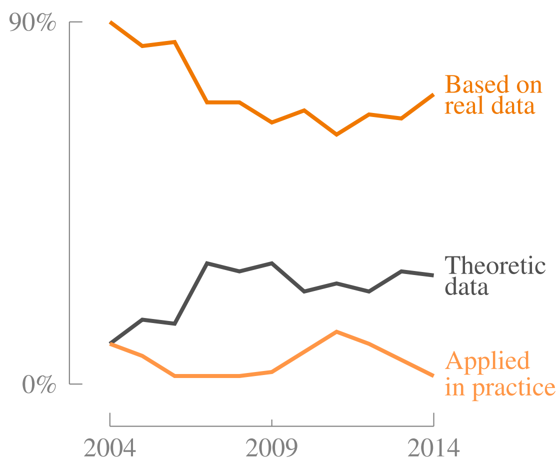
In some cases, MPs are too difficult to solve within a reasonable time limit and therefore heuristics are proposed. Fei et al. [101] use a column generation-based heuristic to solve the patient scheduling problem. In their setting, a column corresponds to a feasible plan representing the assignment of surgical cases to an OR. A genetic algorithm is proposed by Roland et al. [236], which determines the assignment of cases to ORs, planning days and operating time periods.

Some of the articles in the literature use methods that have not been covered in the previous paragraphs. Does et al. [86] use Six Sigma to decrease the tardiness of surgeries, which are performed first on a day. Applied to two hospitals in the Netherlands, substantial savings are achieved and the number of surgeries is increased by 10% without requiring additional resources. Epstein and Dexter [71] introduce a method through which analysts can screen for the economic impact of improving first-case starts. First-case starts are also discussed by Pandit et al. [215].

For future studies, we think that a promising method is simulation-optimization. This method allows to solve complex optimization tasks, while including the complex features of the OR scheduling process.

Also more traditional methods can be used to yield valuable insights. However, the focus should be on making these traditional methods applicable to a broader set of realistic problem settings (e.g., allow multiple sources of variability, broaden the set of supported distributions).

Fig. 2.8 Most data used in the literature are based on real data, however this does not mean that the methods are applied in reality.



2.3.7 Testing phase

Many researchers provide a thorough testing phase in which they illustrate the applicability of their research. Whether applicability points at computational efficiency or at showing to what extent objectives may be realized, a substantial amount of data is desired. From Figure 2.8 and Table 2.9, we notice that most of these data are based on real health care practices. This is noteworthy and results from the improved hospital information systems from which data can be easily extracted.

Investigating the literature, we see that less than 7% of the methods are applied in practice. It seems contradictory that so little research is effectively applied in a domain as practical as OR planning and scheduling.

Unfortunately, simply testing of procedures or tools on real data does not imply that the methods get implemented in real practice. Lagergren [160] indicates that the lack of implementation in the health services seems to have improved considerably. Figure 2.8 shows, however, that only a very small share of the articles report on actual implementation. An exception to this is Wachtel and Dexter [296] who introduce a website, which is used by the hospital under study to decide on the exact times patients have to arrive to their surgery appointment.

The problem tackled by the authors arises from the fact that a case is often started earlier than scheduled, but it cannot be known in advance if it will happen or not. Patient availability must therefore be balanced against patient waiting times and fasting times. Another example is the decision support system of van Essen et al. [93], discussed in Section 2.3.5, for the daily rescheduling problem. Daily applicability is entailed by both methods.

There are problems that have to be solved on a less frequent basis. An example is the application of a case mix model that is applied every year, clearly resulting in a different degree of implementation. A clear comparison of articles on this aspect is hence not straightforward.

Even if the implementation of research can be assumed, authors often provide little detail about the process of implementation. Therefore, we encourage the provision of additional information on the behavioral factors that coincide with the actual implementation. Identifying the causes of failure, or the reasons that lead to success, may be of great value to the research community [47].

A recent example giving insights into these causes is provided by Brailsford et al. [41]. They evaluate the adoption of a particular simulation modeling tool and discuss factors that facilitated or hindered the general adoption of the tool in British health care organizations. Identifying key issues in practice helps the research community to be able to build models that better reflect reality and therefore solve a problem that is closer to the one entailed in practice.

In many articles a problem is defined that is specific to one single hospital and it is unclear whether or to what extent a method is applicable to another setting. In order to justify the generality of their modeling assumptions, Schoenmeyr et al. [246] surveyed several hospitals. Introducing generalizable methods makes it easier to spread and implement good working operations research practices to more than one hospital.

For future work, there could be more research on which planning and scheduling expertise is currently in use in hospitals. Using a survey, Sieber and Leibundgut [252] reported that the state of OR management in Switzerland is far from excellent. A similar more recent exercise for Flemish (Belgium) hospitals is described in Cardoen et al. [52]. We also noticed that few articles build on the results or

Table 2.9 For testing purposes, both theoretic and real data are frequently used.

Based on real data	[2, 3, 5, 9, 11, 14, 15, 16, 17, 18, 21, 22, 23, 27, 28, 29, 37, 38, 39, 43, 44, 45, 48, 49, 53, 56, 60, 63, 64, 67, 68, 69, 70, 72, 75, 76, 77, 78, 80, 83, 85, 90, 92, 94, 101, 102, 103, 104, 105, 107, 108, 110, 111, 113, 114, 118, 119, 122, 124, 125, 127, 128, 134, 135, 137, 141, 142, 144, 148, 150, 151, 152, 156, 157, 159, 166, 170, 173, 174, 175, 176, 182, 183, 185, 187, 188, 189, 190, 192, 196, 197, 200, 201, 203, 204, 205, 206, 209, 210, 211, 213, 214, 216, 217, 218, 219, 220, 221, 222, 225, 227, 230, 235, 236, 241, 242, 246, 247, 249, 250, 258, 262, 265, 266, 267, 270, 271, 272, 275, 276, 278, 280, 281, 282, 286, 288, 290, 292, 295, 296, 302, 304, 306, 309, 310, 312, 314]
Theoretic data	[12, 13, 24, 26, 42, 55, 58, 59, 61, 66, 74, 78, 95, 98, 99, 100, 112, 117, 120, 129, 131, 138, 143, 145, 146, 147, 158, 161, 162, 163, 164, 165, 168, 169, 178, 179, 180, 184, 186, 202, 208, 209, 229, 231, 233, 245, 256, 257, 259, 260, 264, 268, 300, 301, 303, 305, 308, 311, 313]
Applied in practice	[4, 32, 33, 34, 62, 93, 96, 107, 128, 237, 269, 271, 284, 285]

data of other articles. We therefore think that more reproducible research is needed. One way of achieving this is by publishing the data and models that were used. Making the data publicly available (if allowed by the hospital) also allows to determine whether a method is generalizable.

For future research, we suggest to develop guidelines on how, and in what format, scheduling data should be made publicly available. This guideline should also provide a well-defined formal way that allows researchers to easily describe their hospital setting.

2.3.8 Relations between classification fields

So far we looked at classification fields separately. In this section we look at the connections between them (Tables 2.10, 2.11, 2.12).

In Table 2.10 we show how much more likely it is to use stochasticity or a method (e.g., an analytical method) with a certain field B (e.g., deterministic models) compared to field $\neg B$ (e.g., stochastic models). For example, the table shows that analytical and DES models are often used with similar fields. They are both more likely to be used in stochastic environments where capacity questions have to be answered at the discipline level and non-electives are included. While analytical methods are more likely to be applied to isolated problems and tested with theoretic data, DES methods are more often used in an integrated setting and tested with real data. This is understandable as integrating the OR with a up- or downstream unit will generally make analytical models too complex to solve.

Analytical methods seem to lack the flexibility that would allow them to be used in settings where DES models are usable. Moreover, the fact that they are more often tested with theoretic data suggests that articles using analytical methods are more focused on developing the methodology itself rather than on solving an actual real-life problem.

Table 2.10 also shows that MP and improvement heuristics are frequently used with similar fields. Both are often applied to deterministic settings that do not include non-elective arrivals. Whereas MPs are used at all decisions levels, improvement heuristics are usually not applied to capacity-related decisions.

We noticed that improvement heuristic methods are often applied to problems that are computationally too intensive to be solved by an MP. As larger problems tend to represent real-life settings, one might naturally assume that improvement heuristic methods are used for more realistic problems. Interestingly, this might not necessarily be the case as improvement heuristic methods are, as a matter of fact, more often expected to be tested on theoretic data than MPs (Table 2.10).

Similarly, constructive algorithms are mostly tested on theoretic data. This is surprising as one would expect that these algorithms, which allow a high degree of tailoring, would be more often applied to real problem settings.

Table 2.10 also shows results on aspects related to stochasticity. It shows that stochasticity both with regards to arrivals and to surgery durations is mostly applied to discipline and to capacity-related problems. Interestingly, stochasticity is less often used in connection with time assignment problems. This is unexpected as one could argue that in these problems stochasticity is especially important to consider. Furthermore, problems that include non-electives will often

Table 2.10 The likeliness to use stochasticity or a method (Columns) with a specific field (Rows) compared to using it without the specific field ($\neg R$).

Field	Stochasticity: $\frac{P(C\|R)}{P(C\|\neg R)}$		Method: $\frac{P(C\|R)}{P(C\|\neg R)}$					$P(R)$
	Arrivals	Duration	Analytical	DES	MP	Imprv. heur.	Constr. alg.	
Discipline	1.59	1.15	**1.55**	**1.30**	0.95	0.80	0.70	*.15*
Surgeon	0.69	0.60	0.76	0.64	**1.78**	0.16	**0.93**	*.12*
Patient	0.77	0.97	0.63	0.81	1.29	**1.59**	**2.00**	*.74*
Day	0.88	0.68	0.54	0.67	**1.92**	**1.49**	0.95	*.51*
Time	0.31	0.88	0.28	0.84	1.23	**1.61**	**1.54**	*.39*
Room	0.48	0.61	0.29	0.64	**1.89**	**1.54**	1.53	*.53*
Capacity	2.62	1.31	**2.40**	**1.51**	1.07	0.18	0.79	*.33*
Determ.	0	0	0.16	0.12	**1.48**	**1.13**	0.91	*.46*
Stoch. arriv.	0	2.60	**2.76**	**3.44**	0.64	0.67	1.02	*.28*
Stoch. dur.	5.00	0	**3.33**	**4.17**	0.67	0.86	1.36	*.44*
Theor. data	1.23	0.99	1.52	0.32	1.15	**1.69**	**2.24**	*.27*
Real data	0.79	0.99	0.56	**3.04**	**0.85**	0.58	0.48	*.73*
Non-elective	3.51	1.85	**2.89**	**2.05**	0.72	0.51	0.97	*.25*
Isolated	1.05	1.12	**4.48**	0.53	0.93	1.16	**2.27**	*.56*
Integrated	0.96	0.89	0.22	**1.88**	**1.07**	0.86	0.44	*.44*
$P(C)$	*.28*	*.44*	*.15*	*.30*	*.49*	*.23*	*.21*	

Example: the number for Discipline-Arrivals shows that it is 1.59 times more likely to use stochastic arrivals in making decisions on the discipline level than for decisions on other decision levels. For methods, the two largest numbers for each field are shown in bold.

consider both stochastic arrivals and durations. This is positive as non-elective arrival times and their associated added workload are uncertain and are therefore difficult to predict in advance.

Whereas in Table 2.10 the focus is on methods and stochastic aspects, in Tables 2.11 and 2.12 the focus is on PMs and constraints. In Tables 2.11 and 2.12 we use a different measure than in Table 2.10 since we are not interested in the individual importance of a PM/constraint, but in their importance relative to each other. Therefore, we use in Tables 2.11 and 2.12 conditional probabilities, while in Table 2.10 we use ratios of conditional probabilities.

Tables 2.11 and 2.12 show among others that the number of considered PMs is usually higher than the number of included constraints. The largest number of PMs are used in DES models. This is understandable as in DES models the number of PMs does generally not determine the run time of the model. This is in contrast to analytical methods where it can be difficult to include many PMs, which might be a problem in a setting, such as surgery scheduling, to which a large amount of competing PMs and constraints are usually inherent.

Table 2.11 reveals some other interesting connections. For instance, PMs that are mostly used in the DES literature are patient waiting time, overutilization, utilization, throughput and deferral. They are usually used in models that target the discipline level where capacity-related decisions are made and in which real data are used for testing purposes. Understandably, deferral is almost exclusively used in settings where arrivals are modeled stochastically.

It is also noteworthy that utilization and makespan, two measures often used in related operations research fields such as machine scheduling [224], are generally less used in the surgery scheduling literature. Instead, authors seem to prefer to use overtime and, to some extent, undertime. Interestingly, when real data are used for testing purposes, the use of overtime is less probable compared to when theoretic data is used.

Criteria that are used as PMs can generally also be used as constraints. For example, instead of minimizing overtime, a constraint can be defined that limits the allowed overtime to a maximum of 2 hours.

Constraints are included for several reasons. For example, they can be used to represent the limited availability of PACU beds (Up/Downstream), nurses (Personnel) and equipment (Non-personnel). They can also be used to ensure that patients are served before a predefined date (Preferences) or that a minimum number of patients is served by a discipline (Demand).

Understandably, personnel-related constraints are particularly often used in models. These allow to include regulations and rules that are important to the hospital management and staff. The table also shows that preference-related constraints are often applied.

Overall, Tables 2.10, 2.11 and 2.12 can be used to detect (un)common approaches and problem settings. We see that the main problem settings all have been researched to some extent already. One might wonder whether there is anything left to do in OR planning. The fact that practitioners see their problems unsolved, suggests that the job of researchers is not yet done.

Table 2.11 The conditional probabilities of various performance measures given different fields.

Field	Performance measure: $P(C\|R)$									μ_{Count}	$P(R)$
	Patient waiting	Over-util. OR	Under-util. OR	Utiliz. OR	Through-put	Prefe-rence	Finan-cial	Make-span	Defer-ral		
Analytical	**.30**	**.52**	.18	.06	.03	.21	.18	.06	.18	2.30	*.15*
DES	**.48**	**.49**	.18	.40	.29	.09	.08	.09	.25	3.00	*.30*
MP	.24	**.47**	**.25**	.08	.09	.21	.16	.10	.09	2.46	*.49*
Imprv. heur.	**.24**	**.47**	.24	.12	.06	.18	.04	.20	.04	2.12	*.23*
Constr. alg.	**.28**	**.63**	.22	.15	.07	.17	.09	.17	.02	2.37	*.21*
Discipline	.28	**.31**	.16	**.34**	.28	.16	.12	0	.19	2.38	*.15*
Surgeon	.16	**.40**	.12	.16	.08	.08	**.36**	.08	.16	2.32	*.12*
Patient	**.30**	**.52**	.20	.16	.09	.21	.09	.14	.13	2.54	*.74*
Day	.29	**.51**	**.30**	.16	.14	.22	.10	.06	.14	2.59	*.51*
Time	**.26**	**.49**	.12	.16	.15	.13	.06	.24	.11	2.44	*.39*
Room	**.25**	**.53**	.25	.18	.13	.18	.06	.14	.07	2.39	*.53*
Capacity	**.35**	**.39**	.15	.22	.18	.15	.28	.01	.24	2.76	*.33*
Determ.	.17	**.36**	.21	.12	.08	**.24**	.10	.18	.05	2.09	*.46*
Stoch. arriv.	**.47**	**.45**	.15	.30	.18	.15	.17	.02	.32	2.88	*.28*
Stoch. dur.	**.44**	**.59**	.21	.23	.18	.12	.10	.05	.17	2.89	*.44*
Theor. data	**.29**	**.58**	.27	.07	.03	.17	.14	.14	.10	2.42	*.27*
Real data	**.28**	**.40**	.18	.23	.16	.18	.11	.10	.13	2.44	*.73*
Non-elective	**.42**	**.58**	.17	.26	.13	.19	.11	.04	.26	2.77	*.25*
Isolated	**.28**	**.50**	.21	.19	.08	.18	.12	.08	.12	2.20	*.56*
Integrated	**.29**	**.36**	.19	.17	.19	.18	.10	.16	.12	2.69	*.44*
$P(C)$	*28*	*.44*	*.20*	*.18*	*.12*	*.18*	*.12*	*.11*	*.12*		

Example: the value 0.30 represents the conditional probability of the occurrence of patient waiting time given an analytical method. The one but last column shows the average number of PMs used with the specific field. For example, on average 2.30 PMs are used with an analytical model. The two largest numbers for each field are shown in bold.

In particular, we suggest two directions for future research. First, there are still new topics to be further explored (highlighted in the final paragraphs of Section 2.3.1 - 2.3.7). Second, there are already researched problems of which the solutions are not used by stakeholders and therefore need to be revisited. In order to get solutions more applied by stakeholders, we define the following three actions: authors need to (1) decide whether they want to address researchers or practitioners and build up their article according to this decision, (2) select PMs that fit their (hospital) setting, not their used model and (3) clearly specify setting- and method-specific assumptions in the text. These three points are the main topics discussed in Section 6.

Table 2.12 The conditional probabilities of various constraints given different fields.

Field	Constraint: $P(C\|R)$					μ_{Count}	$P(R)$
	Up/Downstr.	Personnel	Non-person.	Preferences	Demand		
Analytical	.03	**.24**	.15	**.27**	.09	.88	*.15*
DES	**.18**	**.45**	.09	.18	.15	1.51	*.30*
MP	.26	**.74**	**.36**	.36	.31	2.75	*.49*
Imprv. heur.	.14	**.55**	.14	**.35**	.12	1.65	*.23*
Constr. algo.	.13	**.63**	.26	**.30**	.13	1.78	*.21*
Discipline	.06	**.47**	.22	.28	**.50**	2.09	*.15*
Surgeon	.20	**.76**	.16	.16	**.56**	2.56	*.12*
Patient	.20	**.56**	.24	**.35**	.14	2.03	*.74*
Day	.17	**.71**	.30	**.35**	.32	2.57	*.51*
Time	.25	**.64**	.28	**.40**	.16	2.33	*.39*
Room	.19	**.71**	.29	**.38**	.23	2.40	*.53*
Capacity	.17	**.54**	.21	.17	**.35**	2.06	*.33*
Determ.	.25	**.65**	.31	**.41**	.21	2.59	*.46*
Stoch. arriv.	.12	**.48**	.17	.17	**.20**	1.50	*.28*
Stoch. dur.	.14	**.48**	.14	**.20**	.18	1.48	*.44*
Theor. data	.14	**.54**	.27	**.29**	.14	1.69	*.27*
Real data	.21	**.56**	.20	**.29**	.22	2.08	*.73*
Non-elective	.17	**.40**	.13	**.26**	.13	1.57	*.25*
Isolated	.02	**.53**	.18	**.32**	.13	1.41	*.56*
Integrated	**.41**	**.56**	.27	.24	.28	2.66	*.44*
$P(C)$	*.19*	*.55*	*.22*	*.29*	*.20*		

2.4 Conclusion

In this chapter we classified and discussed the OR planning and scheduling literature. We classified the literature with regard to the patient type, the different performance measures, the decisions that have to be made, the integration of up- and downstream units of the OR, the incorporation of uncertainty, the operations research methodology and the testing phase. The introduced classification tables help the reader to quickly identify relevant articles.

We also looked at trends for the last ten years and determined that the amount of published technical articles has been increasing significantly in the recent ten years. We also analyzed of the connections between the classification fields in order to show which methods, PMs and constraints are commonly combined and which are not. At the end of each section we highlighted topics for future research.

Chapter 3

Hospital setting and model

This chapter contains a description of the hospital setting and the details on the simulation model. We describe the processed hospital data, patient attributes, the MSS, the non-elective allocation schema and the way patients are rescheduled. In order to validate the simulation model, we compare OR-related performance measures to actual measurement data from the University Hospital Leuven.

3.1 Hospital setting

The hospital provided us with patient scheduling records of the complete years 2012/2013 and helped us to correctly interpret those records. The data was grouped into three datasets: Patient trajectory, OR information and Patient planning related data. The first data set, Patient trajectory, contains 922.900 entries and 23 attributes. The dataset contains, on the one hand, the date and time patients were transferred to different facilities in the hospital and, on the other hand, general information concerning the patient such as the arrival date, the ID of the surgeon in charge and whether the patient arrived as an emergency. The second dataset, OR information, contains 34.288 entries and 14 attributes.

The dataset includes surgery specific information, such as the ID of the planned surgical intervention, the estimated surgery duration and the DT of the patient. The last dataset, Patient planning, contains 56.912 entries and 16 attributes. The data contains surgery scheduling related records, such as the surgery's planned date and OR. A new entry is created in the table anytime the planner changes the scheduling information of a patient.

Before merging the three datasets, each of them was preprocessed individually. Preprocessing involved, for example, combining entries in a dataset that relate to the same surgery. The datasets were merged in a way that all the information attached to one surgery became easily accessible and processable. Some entries were lacking important attributes, such as the arrival data and were therefore removed from the dataset. After removing all uncompleted entries 32.042 patients remained from 13 disciplines.

3.1.1 Patient arrivals

The arrival time of elective patients is the time point when their surgeon determines the need for surgery. This generally happens on weekdays at any point during the daytime. The arrival time of non-elective patients represents the time point when they are physically registered at the hospital. This can happen at any day and at any hour.

We model elective patient arrivals on the basis of a statistic that is based on rates, e.g., 5 CAH patients request surgery on a Monday (Fig. 3.1). For non-electives, a statistic is used that is based on inter-arrival times. This defines an exact time instance, e.g., a non-elective CAH patient arrives Monday at 2.21 pm.

Table 3.1 shows that the mean number of electives requesting surgery weekly is 240.35 with a standard deviation of 32.53 (column $\cup$ in the table). The average number of arrivals for an elective discipline is 18.49 with a standard deviation of 5.6 (column μ in the table).

As Table 3.1 shows, patient arrival numbers are highly variable. This is true for week to week (e.g., first week to second week of the year), day to day (e.g., Monday to Tuesday) and weekday to weekday (e.g., Tuesday to Tuesday) based

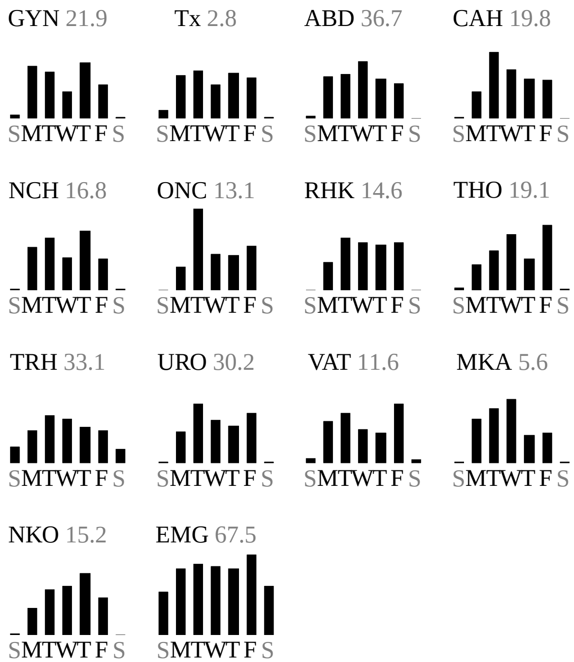


Fig. 3.1 The average arrival rate of the 13 elective disciplines and non-electives. The number in gray represents the average weekly number of arrivals. The height of the column represents the percentage of patients that arrived on that day from the weekly volume. The figure is based on data covering the entire years 2012-2013.

comparisons. It is especially surprising that the weekday to weekday variation of patient arrivals is high. This might be counterintuitive as, given that surgeons have consultation times on a fairly regular basis (e.g., every Monday), one could assume that patient arrival numbers for the same weekday are more stable.

Interestingly, the week to week arrival variability differs strongly between disciplines. For MKA (and Tx) it is very high in relation to the mean, resulting in a large CV. Consequently one might wonder whether it is possible to provide timely service to MKA patients. Fortunately, MKA patients are generally not urgent (DT score of 0.31, Table 1.2). This allows to spread out arrivals from weeks with high loads to weeks with lower loads. The same could not be done by TRH as most of their patients must be served within 1 week. Fortunately, TRH has one of the most stable patient inflows and will therefore less frequently encounter weeks with very high loads.

One could assume that in reality disciplines with a patient mix that contains higher urgency patients or a larger arrival variability would generally provide less timely service to their patients when compared to the rest of the disciplines. Interestingly, we did not find any indications in the data that would support this theory.

In case a discipline covers a large population of DT 4 patients, not only the weekly, but also the daily arrival variability is important. Consequently, in the model, both discipline-dependent weekly and also daily arrival variability needs to match reality. We ensure this by generating patients in two steps. In the first step, we determine for each discipline the total number of weekly arrivals. In the second step, the number of arrivals for each weekday is determined (Monday to Friday). This is done by selecting a realization of a week from a pool. The weeks in the pool were pre-generated using the empirical distributions observed in reality.

As a consequence of this two-step procedure, for all disciplines, the difference between the model and the reality of arrival means and standard deviations are minimal (Table 3.1). A difference is present only if the union of all patients is considered. This difference can be explained by holidays. In reality, holidays in a week result in lower arrival numbers for all disciplines, i.e., weekly arrival numbers for disciplines correlate. In the model, holidays affect each discipline independently, therefore weekly arrival numbers for disciplines do not correlate. With regards to the results, this discrepancy will not matter because patients of different disciplines are scheduled into their own OR capacity, i.e., while the individual arrival variability of each discipline is important, this is less the case for the combined one.

In Table 3.1 it is interesting to observe that arrival means are generally not equal to their variance. This is the case for most disciplines, for the averages across disciplines (denoted in the table by μ) and for combined elective arrivals (denoted in the table by $\cup$). Furthermore, this is true for weekly arrival numbers, daily arrival numbers and weekday specific arrival numbers (Fig. 3.1). The fact that the arrival means and variances are not equal means that the arrivals, contrary to what is sometimes assumed in the literature, do not follow the Poisson distribution. Interestingly, this is even true for non-elective weekly arrival numbers. The two elective categories that seem to be exceptions are Tx and NKO.

A factor that has an influence on arrival variance are holidays. The number of arrivals on holidays is lower than on normal days, but is by far not zero, i.e., patients are also scheduled for surgery on holidays. It is important to note that excluding holidays will decrease the arrival variability only to a limited extent.

3.1.2 Non-electives

Every week around 70 non-elective patients, using around 160 hours of OR time (Fig. 3.5), get surgery at the hospital. This means that, if scheduled into regular OR time, they would occupy 3 to 4 ORs a day. This is a large number which explains their fundamental impact on the hospital's OR department. In order to realistically model this impact, we analyzed both their arrival patterns and the discipline-dependent way they are allocated to ORs (Fig. 3.2).

Non-electives arrive with the highest rate during daytime on weekdays. We call those time intervals high impact periods as this is also the time when non-electives have the largest impact on the elective schedule.

In the DES model, we explicitly model high impact periods, i.e., non-elective inter-arrival times will depend on the day of the week (Fig. 3.1) and the time period of the day. There are two time periods, (1) daytime is between 6 am and 10 pm and (2) nighttime is between 10 pm and 6 am. Arrival ratios will be around 3.4 times higher during daytime than during nighttime.

Another important component of the model is the discipline-dependent non-elective OR allocation schema. As shown in Figure 3.2, during high impact periods non-electives of all DT categories are generally served in an OR of the corresponding discipline. For example, an open wound patient brought to the

Table 3.1 The arrival statistics measured at the hospital compares well to the arrival statistics produced by the model (Δ values are small).

Arrivals		Non-Elective	Elective														
			GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	μ	∪
Weekly																	
real	μ	**67.6**	**21.9**	**2.8**	**36.7**	**19.8**	**16.8**	**13.1**	**14.6**	**19.1**	**33.1**	**30.2**	**11.6**	**5.6**	**15.2**	**18.5**	**240**
	var	118	38.8	2.9	84.7	50.1	36.7	15.9	39.1	39.4	51.8	46.1	20.8	9.9	18.5	35.0	1058
	σ	*10.9*	*6.2*	*1.7*	*9.2*	*7.1*	*6.1*	*4.0*	*6.3*	*6.3*	*7.2*	*6.8*	*4.6*	*3.1*	*4.3*	*5.6*	*32.5*
	CV	.16	.28	.61	.25	.36	.36	.30	.43	.33	.22	.22	.39	.56	.28	.35	.14
model	μ	**67.4**	**22.1**	**2.7**	**36.6**	**19.6**	**16.8**	**13.4**	**14.9**	**19.2**	**32.8**	**30.4**	**11.6**	**5.7**	**15.0**	**18.5**	**241**
	var	75.1	40.5	2.7	85.4	45.4	38.4	16.5	41.6	39.7	51.9	44.8	21.8	9.8	18.0	35.1	481
	σ	*8.7*	*6.4*	*1.6*	*9.2*	*6.7*	*6.2*	*4.1*	*6.5*	*6.3*	*7.2*	*6.7*	*4.7*	*3.1*	*4.2*	*5.6*	*21.9*
	CV	.13	.29	.60	.25	.34	.37	.30	.43	.33	.22	.22	.40	.55	.28	.35	.09
Δ	μ	**.20**	**-.16**	**.07**	**.04**	**.14**	**-.06**	**-.27**	**-.29**	**-.06**	**.23**	**-.21**	**-.07**	**-.11**	**.18**	**-.04**	**-.58**
	var	43.1	-1.7	.20	-.68	4.6	-1.7	-.59	-2.5	-.24	-.08	1.3	-1.1	.06	.47	-.14	577
	σ	*2.2*	*-.13*	*.06*	*-.04*	*.34*	*-.14*	*-.07*	*-.20*	*-.02*	*-.01*	*.09*	*-.11*	*.01*	*.05*	*-.01*	*10.6*
	CV	.03	.00	.01	.00	.01	-.01	.00	.00	.00	.00	.00	-.01	.01	.00	.00	.04
Daily																	
real	μ	**10.7**	**4.3**	**.53**	**7.2**	**3.9**	**3.3**	**2.6**	**2.9**	**3.8**	**5.7**	**5.9**	**2.2**	**1.1**	**3.0**	**3.6**	**46.5**
	var	11.6	6.1	.59	10.8	6.7	4.6	4.3	4.3	6.1	7.9	8.3	2.9	1.4	3.4	5.2	122
	σ	*3.4*	*2.5*	*.77*	*3.3*	*2.6*	*2.2*	*2.1*	*2.1*	*2.5*	*2.8*	*2.9*	*1.7*	*1.2*	*1.8*	*2.2*	*11.0*
	CV	.32	.58	1.5	.46	.66	.65	.79	.71	.66	.49	.49	.76	1.1	.61	.72	.24
model	μ	**10.5**	**4.3**	**.51**	**7.2**	**3.9**	**3.3**	**2.7**	**3.0**	**3.8**	**5.7**	**6.0**	**2.2**	**1.1**	**3.0**	**3.6**	**46.7**
	var	11.0	6.6	.56	11.6	7.5	5.3	4.5	5.1	6.8	8.0	8.3	3.3	1.5	3.4	5.6	97.3
	σ	*3.3*	*2.6*	*.75*	*3.4*	*2.7*	*2.3*	*2.1*	*2.3*	*2.6*	*2.8*	*2.9*	*1.8*	*1.2*	*1.9*	*2.3*	*9.9*
	CV	.31	.59	1.5	.47	.70	.69	.80	.76	.69	.50	.48	.81	1.1	.62	.74	.21
Δ	μ	**.16**	**-.03**	**.02**	**.01**	**.02**	**-.02**	**-.06**	**-.06**	**-.01**	**.03**	**-.05**	**-.02**	**-.02**	**.03**	**-.01**	**-.15**
	var	.65	-.46	.03	-.80	-.85	-.66	-.27	-.87	-.63	-.09	-.02	-.39	-.12	-.01	-.40	24.7
	σ	*.10*	*-.09*	*.02*	*-.12*	*-.16*	*-.15*	*-.06*	*-.20*	*-.12*	*-.02*	*.00*	*-.11*	*-.05*	*.00*	*-.08*	*1.2*
	CV	.00	-.02	-.01	-.02	-.04	-.04	-.01	-.05	-.03	-.01	.00	-.04	-.02	-.01	-.02	.03
Weekday (Tuesday)																	
real	μ	**10.8**	**4.6**	**.60**	**7.3**	**5.9**	**4.0**	**4.8**	**3.5**	**3.4**	**7.0**	**7.9**	**2.6**	**1.4**	**3.1**	**4.3**	**56.1**
	var	12.5	5.8	.63	9.1	8.8	4.6	5.4	4.2	3.6	8.4	7.9	2.9	1.7	2.8	5.1	106
	σ	*3.5*	*2.4*	*.79*	*3.0*	*3.0*	*2.1*	*2.3*	*2.0*	*1.9*	*2.9*	*2.8*	*1.7*	*1.3*	*1.7*	*2.2*	*10.3*
	CV	.33	.52	1.3	.41	.50	.54	.48	.59	.56	.41	.35	.66	.96	.54	.61	.18
model	μ	**10.6**	**4.7**	**.62**	**7.2**	**6.0**	**4.1**	**4.9**	**3.5**	**3.3**	**7.0**	**7.7**	**2.5**	**1.4**	**2.9**	**4.3**	**55.9**
	var	10.2	6.2	.69	10.4	11.4	5.4	5.3	4.3	3.9	8.2	8.5	3.3	1.8	2.4	5.5	76.7
	σ	*3.2*	*2.5*	*.83*	*3.2*	*3.4*	*2.3*	*2.3*	*2.1*	*2.0*	*2.9*	*2.9*	*1.8*	*1.3*	*1.6*	*2.2*	*8.8*
	CV	.30	.53	1.4	.45	.56	.56	.47	.59	.61	.41	.38	.71	.96	.54	.62	.16
Δ	μ	**.19**	**-.10**	**-.02**	**.09**	**-.06**	**-.15**	**-.08**	**-.03**	**.14**	**-.01**	**.22**	**.03**	**-.01**	**.18**	**.01**	**.20**
	var	2.3	-.42	-.06	-1.3	-2.6	-.76	.08	-.12	-.32	.22	-.55	-.42	-.05	.39	-.45	29.1
	σ	*.34*	*-.09*	*-.03*	*-.21*	*-.41*	*-.17*	*.02*	*-.03*	*-.08*	*.04*	*-.10*	*-.12*	*-.02*	*.12*	*-.08*	*1.5*
	CV	.03	-.01	-.01	-.03	-.06	-.02	.01	.00	-.05	.01	-.02	-.05	-.01	.01	-.02	.03

The table also shows that arrivals generally do not follow the Poisson distribution (mean and variance are not equal). Weekly means are calculated on the basis of the 104 weeks of the years 2012 and 2013. Daily means are calculated on the basis of 520 days whereas the values of weekdays are calculated on the basis of the corresponding 104 weekdays (e.g., all Tuesdays in 2012 and 2013). In the table, as an example, only Tuesday is shown. The mean value of all elective disciplines is denoted by 'μ', whereas the value considering electives in general is denoted by '∪'.

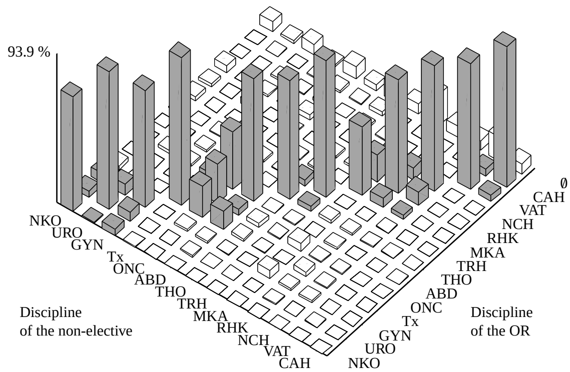


Fig. 3.2 The comparison between non-elective disciplines and the discipline of the OR the surgery was carried out shows that the two usually correspond. Non-electives are generally assigned to an OR that serves electives of that discipline (diagonal columns). Occasionally, non-electives can be served in ORs that are not assigned to any discipline (marked with ∅). Disciplines in the figure are grouped on the basis of their cluster (in gray). These clusters contain 4-6 ORs that form a physical unit. The figure is based on data covering the entire years 2012-2013.

hospital is generally assigned to an OR that is occupied by electives from TRH. Non-elective ONC patients are the only exception to this rule as they are frequently served in ORs allocated to ABD or Tx. These three disciplines, however, belong to the same cluster.

In the DES model, during high impact periods non-electives can generally only enter an OR that serves patients of their discipline. An exception is made for MKA, ONC and Tx as there will be weekdays on which they have no OR assigned to them. In those cases MKA patients are assigned to empty ORs. Tx can always occupy OR 7 even if the OR was originally closed on the day, whereas ONC patients can enter ORs of ABD and Tx. Those exceptions are based on our findings in the hospital data and thus imitate the real practice.

In the model, we also distinguish between DT category 1 and DT categories 2 and 3. DT category 1 patients have to be served immediately (Fig. 3.3) and are assigned to the next possible suitable open OR serving their discipline [167]. Contrarily, DT category 2 and 3 patients will be added to the end of the schedule. This is also often happening in reality and serves the interest of the surgeons as it allows them to finish all their electives before starting any non-elective.

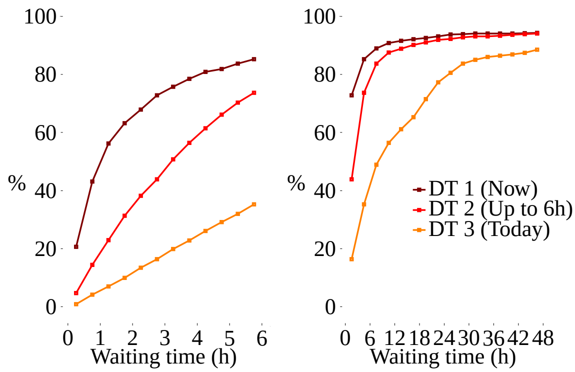


Fig. 3.3 The cumulative distribution function of non-elective (direct) waiting time. Around 70% of DT 1 patients are served within 3 hours after arriving to the hospital. The figure is based on data covering the entire years 2012-2013.

3.1.3 Surgery duration

The surgery duration of a patient is defined as the time that elapsed between the moment the patient is rolled into the OR and the time when the patient leaves the OR (Fig. 3.4). It does not include cleaning time. If the patient is already present in the OR, the surgery duration includes the setup time. Generally, if the setup time is specific to the patient, then it is included into the surgery duration.

The estimated surgery duration (Table 3.2), suggested to the surgeon, is based on the mean of the realized surgery durations of previous similar OR sessions. The surgeon can then accept or overrule this value.

Each discipline performs different types of surgeries. Each of those surgery types is assigned a unique identifier that generally contains an ICD-9 code and a local component. ICD-9 codes by themselves can be too restrictive to accurately describe a procedure and thus need this additional local component. Surgeries with the same identifier will represent similar procedures and will consequently have a similar estimated length.

As Table 3.3 shows, the log-logistic distribution provides, from all tested parametric distributions, the best fit on surgery types. The log-normal distribution is sometimes used in the literature as it provides a better fit than the normal

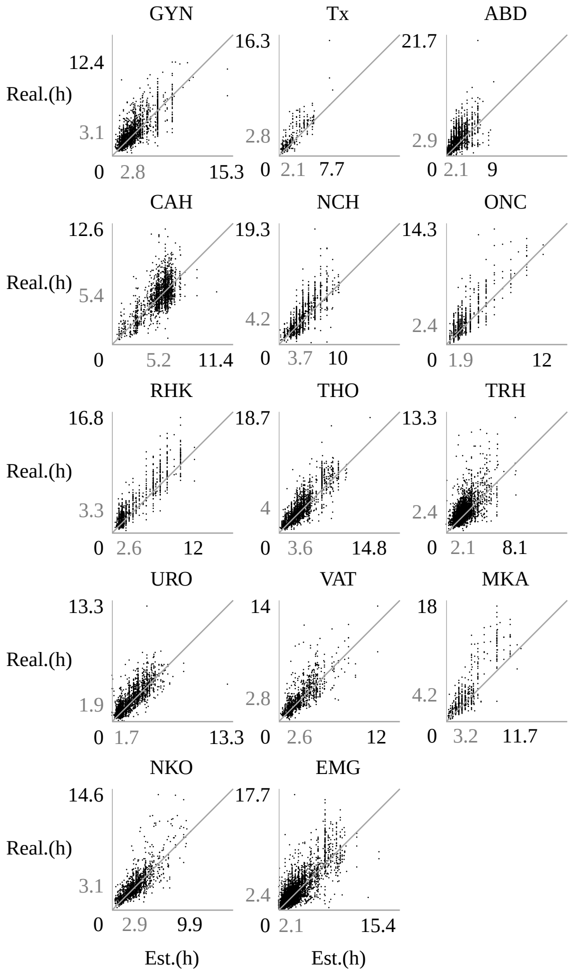


Fig. 3.4 Comparison of the estimated (x axis) and realized (y axis) surgery durations. Points on the same vertical line are either surgeries of the same type, surgery types with the same duration or they are discrete values that were selected manually by surgeons (e.g., 1 hour and 30 min). The numbers in gray represent the mean values. Higher resolution graphs are included in Appendix A.

Table 3.2 Comparing realized and estimated (planned) surgery durations (hours) shows that surgery durations are systematically underestimated (i.e., realized surgery durations are usually longer than estimated surgery durations).

		Non-Elective	Elective														
			GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	μ	$\cup$
Realized	μ	**2.4**	**3.1**	**2.8**	**2.9**	**5.4**	**4.3**	**2.4**	**3.3**	**4.0**	**2.5**	**1.9**	**2.8**	**4.2**	**3.1**	**3.3**	**3.2**
	σ	*2.0*	*1.8*	*2.0*	*1.6*	*1.7*	*2.4*	*1.9*	*2.8*	*2.5*	*1.3*	*1.4*	*1.8*	*3.4*	*1.8*	*2.0*	*2.1*
	CV	.86	.56	.70	.54	.32	.55	.78	.86	.62	.53	.75	.63	.80	.57	.63	.67
Est.	μ	**2.1**	**2.8**	**2.1**	**2.1**	**5.2**	**3.8**	**1.9**	**2.6**	**3.6**	**2.1**	**1.7**	**2.6**	**3.2**	**2.9**	**2.8**	**2.7**
	std	1.8	1.6	1.4	1.1	1.3	1.7	1.5	2.4	2.1	.89	1.2	1.5	2.1	1.5	1.6	1.8
	CV	.84	.57	.68	.52	.25	.45	.77	.93	.59	.42	.71	.59	.67	.49	.59	.65
Δ	μ	**.29**	**.30**	**.70**	**.78**	**.16**	**.50**	**.48**	**.69**	**.46**	**.33**	**.18**	**.20**	**1.0**	**.15**	**.46**	**.42**
	std	1.2	1.2	1.2	1.2	1.3	1.4	1.0	1.1	1.3	1.0	.77	1.1	1.9	1.1	1.2	1.2
	CV	4.0	4.0	1.7	1.5	8.0	2.9	2.1	1.6	2.9	3.1	4.2	5.7	1.9	7.3	3.6	2.8

The high standard deviation of the error means that surgery durations are often misestimated by a large number of hours. This is a problem as large estimation errors lead to OR overtime, case cancellations and generally decreased efficiency of OR resources [289]. As shown in Table B.1 (Appendix B), for most disciplines it is only a few surgeons that are responsible for the majority of surgery duration underestimation.

distribution [263]. Also in our setting, the log-normal distribution clearly dominates the normal distribution. However, importantly, the log-logistic distribution outperforms both of them. Firstly, the log-logistic distribution fits all of the surgery types, whereas the log-normal distribution fits 97.7% of the surgery types. Additionally, based on the AIC criterion, the log-logistic distribution provides, amongst the tested distributions, the best fit in 31.8 % of the cases. For the log-normal distribution this is true for 2.7 % of the cases whereas the normal distribution never provides the best fit.

Despite the fact that surgery types seem easy to work with, there is a factor that prevents their use. We will often lack a sufficiently large sample size to reliably estimate the parameters of a distribution. The problem would remain if we would analyze more than two years of data. A larger total sample size would likely include new unseen surgery types which might again have a low sample count.

Because of the aforementioned factors, we model surgeries on a higher level, namely on the level of the discipline (Fig. 3.4). This avoids the problem of low counts, but unfortunately also introduces a new problem, namely multimodality. This is the case when disciplines cover several surgery types, which typi-

Table 3.3 The Kolmogorov-Smirnov test is used to compare the sample surgery durations per type with the referenced probability distributions.

Distribution	% of types with good fit	% of types with best fit (AIC)
Log-logistic	100%	31.8%
Logistic	98.5%	3.5%
Log-normal	97.7%	2.7%
Gamma	96.7%	5%
Birnbaumsaunders	96.3%	8%
Inverse gaussian	96.2%	26.6%
Nakagami	94.5%	5.3%
Weibull	92.8%	9.7%
Rician	92.2%	1.3%
Normal	91%	0%
Extreme value	75.2%	4.3%
Rayleigh	41.1%	1.8%

For each surgery type, a ranking is created using the Akaike information criterion (AIC). Only those surgery types that are performed at least 10 times during the years 2012-2013 were included into the analysis. This covers 78% of all surgeries.

cally have a different mean duration. Unimodal parametric distributions (such as described in Table 3.3) do not work well on multimodal data. Methods that do work are based on a kernel density estimator (KDE) or a Gaussian mixture model (GMM).

In Table 3.4 we compare the goodness of fit of a few bivariate models on the data. The first distribution in the table is a purely parametric model, the bivariate GMM. The remaining models are based on the theory of copulas.

Both GMMs and copula-based models have their benefits and drawbacks. A GMM assumes that all the data points are generated from a mixture of a finite number of Gaussian distributions. In reality, this assumption might not be true for surgery durations.

Copulas are not constrained to distributions with Gaussian mixes. Copulas provide a way to describe joint distributions by separating the estimation of the marginal distributions of the random variable from the dependencies between them. Unfortunately, copulas such as the Gaussian- or (Student) t-copula come with their own set of restrictions as they can perform worse on multimodal data [273].

In order to model realized and estimated surgery durations at the discipline level, a model is needed that can handle multimodality and is flexible with regards to the assumptions made on the underlying distribution. Such a model was developed by Tewari et al. [273] and is a combination of GMMs and copulas using a class of functions called Gaussian Mixture Copula (GMC) functions. In Table 3.4 we compare such a GMC-based model with a bivariate GMM, a Gaussian copula and a Student-t copula.

From Table 3.4 we see that the bivariate GMM performs well with regards to some disciplines. For the disciplines where the bivariate fit is bad, the marginal fit on estimated durations is bad as well. The bad fit is likely a consequence of the fact that estimates can have a discrete component.

On the contrary, as Table 3.4 shows, copula models can provide a good fit on the marginals, but do not perform well with regards to the bivariate fit. This shows that the method fails to correctly capture the connection between the realized and the estimated durations. More specifically, both the Gaussian and the t-copulas seem to fail because of the multimodal aspect of the data.

A method that provides a good fit on both the bivariate distribution and on the marginals is the GMC model. As generally with copula methods, it is also in this case of critical importance to choose suitable marginal distributions. For example, choosing a log-normal marginal distribution for the realized durations clearly yields a bad fit (Table 3.4). Two marginals that work well are the univariate GMM and the KDE. The KDE we found to work well is the fixed bandwidth method described by Shimazaki and Shinomoto [248].

From our analysis, we conclude that in order to realistically model surgery durations in an inpatient setting the following rules are important. If only realized durations are modeled on the level of surgery types, the log-logistic distribution should be used. If realized durations are modeled on an aggregated level (e.g., discipline), we advice to use a GMM as this is a simple parametric distribution that provides a good fit. If both realized and estimated durations are modeled, then either a fully empirical distribution should be chosen or a bivariate copula model that is able to handle multimodality (e.g., GMC). The marginal distribution of the copula model should be based on a GMM or a KDE. If estimated durations contain a pronounced discrete component, the corresponding marginal distribution should be based on a KDE.

In the simulation model, we use the described GMC model with a univariate GMM for realized and a KDE method for estimated marginals. We did not use a purely empirical model as the duration generator for some of the disciplines with lower sample counts (e.g., Tx) would produce reoccurring duration values.

3.1.4 Capacity allocation

In the literature, the OR planning process is commonly divided into three levels: strategic, tactical and operational [25]. At the strategic level, a certain amount of OR capacity is allocated to each discipline. This relates to the patient case mix as the hospital decides for each discipline on the number of future patients it wants to serve. At the tactical level, the MSS is created, this is a 1- or 2-week cyclic plan where to each weekday (or half a day) and OR a specific discipline or surgeon is assigned. At the operational level, surgeons assign patients to their own OR sessions. There are hospitals where the ORs are planned differently, but generally a schema similar to the one described is followed.

There are many criteria that can guide the creation of the MSS. Typically, the average arrival caseload and its variability are factors that are considered [136]. Additional factors can relate to tradition, i.e., if a discipline generally received a lot of capacity, they might also get more capacity in the future.

In the simulation model the MSS is predetermined and therefore static. This also means that the capacities assigned to each discipline are fixed for each week. The fixed weekly capacities we use in the model are equal to the average capacities of the University Hospital Leuven's final MSS. For example, if in reality on average 5.75 ORs a week are used by NKO, then in the simulation model an MSS with four cycles is used where NKO in one week is assigned five ORs while in the other three weeks six ORs are assigned to NKO.

There will be slack capacity in the system as the supply of OR time is larger than the demand. The slack capacity shown in Table 3.6 is based on duration

Table 3.4 The goodness of fit tests for various bivariate models (realized and estimated surgery duration pairs) and their marginals (only realized / estimated durations) shows that only the GMCM copula can provide a good fit on the joint distribution.

	Non-Elective	Elective												
Method		GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO
GMM-Biv														
joint	<.001	<.001	**.73**	.02	**.91**	**.19**	.02	<.001	**.6**	**.9**	**.79**	**.92**	**.81**	**.06**
real.	**.26**	**.29**	**.53**	**.92**	**.86**	**.55**	**.22**	.02	**.19**	**.89**	**.13**	**.91**	**.53**	**.75**
est.	<.001	<.001	**.09**	<.001	**.5**	<.001	<.001	<.001	.03	**.07**	**.18**	**.31**	**.11**	**.83**
Gauss.														
joint	<.001	.05	**.11**	.02	<.001	<.001	<.001	<.001	<.001	0	<.001	<.001	<.001	**.37**
GMM-real.	**.47**	**.15**	**.99**	**.22**	**.51**	**.65**	**.94**	**.47**	**.2**	**.08**	**.51**	**.46**	**.73**	**.07**
KDE-est.	**.64**	**.76**	**.19**	**.52**	**.83**	**.72**	**.52**	**.86**	**.95**	**.72**	**.72**	**.3**	**.54**	**.3**
Stud.-t														
joint	<.001	**.08**	**.06**	**.16**	<.001	<.001	0	<.001	<.001	**.4**	<.001	**.43**	<.001	**.51**
GMM-real.	**.53**	**.53**	**.86**	**.51**	**.12**	**.26**	**.95**	**.53**	**.26**	**.47**	**.48**	**.48**	**.79**	**.26**
KDE-est.	**.83**	**.58**	**.23**	**.88**	**.78**	**.93**	**.43**	**.56**	**.98**	**.58**	**.43**	**.34**	**.48**	**.46**
GMCM														
joint	<.001	**.2**	**.1**	<.001	<.001	.04	.03	<.001	<.001	<.001	<.001	**.13**	<.001	**4**
logN-real.	<.001	**.06**	**.12**	<.001	<.001	**.15**	<.001	<.001	<.001	<.001	<.001	**.14**	<.001	<.001
KDE-est.	**.64**	**.74**	**.62**	**.88**	**.97**	**.88**	**.94**	**.87**	**.82**	**.75**	**.23**	**.38**	**.68**	**.84**
GMCM														
joint	<.001	**.05**	**.87**	**.13**	**.14**	<.001	.01	**.31**	**.98**	**.29**	**.65**	**.38**	**.86**	**.32**
GMM-real.	**.99**	**.29**	**.68**	**.95**	**.33**	**.87**	**.63**	**.46**	**.75**	**.38**	**.99**	**.49**	**.8**	**.71**
GMM-est.	<.001	0	.01	<.001	<.001	**.35**	**.11**	<.001	.04	**.05**	**.64**	**.45**	**.52**	**.27**
GMCM														
joint	**.37**	**.95**	**.9**	**.12**	**.56**	**.33**	**.91**	**.23**	**.93**	**.19**	**.43**	**.11**	**.25**	**.55**
KDE-real.	**.95**	**.99**	**.71**	**.31**	**.88**	**.41**	**1**	**.58**	**.22**	**.41**	**.92**	**.67**	**.56**	**.79**
KDE-est.	**.51**	**.84**	**.76**	**.24**	**.66**	**.93**	**.86**	**.45**	**.84**	**.56**	**.94**	**.62**	**.88**	**.88**
GMCM														
joint	**.4**	**.75**	**.75**	**.95**	**.81**	**.41**	**.31**	**.92**	**.48**	**.69**	**.11**	**.6**	**.73**	**.3**
GMM-real.	**.59**	**.79**	**.91**	**.36**	**.28**	**.86**	**.68**	**.8**	**.8**	**.72**	**.58**	**.85**	**.9**	**.91**
KDE-est.	**.98**	**.56**	**.35**	**.61**	**.8**	**.61**	**.93**	**.58**	**.56**	**.78**	**.09**	**.58**	**.8**	**.85**

As the Kolmogorov-Smirnov test is only applicable to continuous distributions, we use a χ^2 test. For the joint, bivariate distribution, a two-sample two-dimensional χ^2 test is used whereas for the marginals (realized and estimated) a two-sample one-dimensional χ^2 test is used. The bins in the two-dimensional case are based on a 10*10 grid of bins, whereas in the one-dimensional case on 10 bins. Bins with a count lower than 5 are merged with neighboring bins. Distributions where all p-values are in bold (> 0.05) provide a good fit with the data and can therefore be used in the model.

Table 3.5 The MSS used in the simulation model is a combination of template week A and week B.

Week A / week B Room	MON	TUE	WED	THU	FRI
			Cluster A		
A1	URO	URO	URO	URO	URO
A2	NKO	NKO	NKO	NKO	NKO
A3	GYN	GYN	GYN	GYN	GYN
A4	GYN	GYN	[]	URO	NKO / []
			Cluster B		
B1	ABD	ABD	ONC	ABD	ABD
B2	ABD	ABD	ABD	ABD	ABD
B3	[]	Tx	[]	[]	Tx
B4	ONC	[]	[]	ABD / []	ABD
			Cluster C		
C1	THO	THO	THO	THO	THO
C2	TRH	THO	THO	THO	TRH
C3	TRH	TRH	TRH	TRH	TRH
C4	TRH / []	THO / []	TRH	URO / []	ONC
			Cluster D		
D1	NCH	NCH	NCH	NCH	NCH
D2	RHK	RHK	RHK	RHK	RHK
D3	MKA / []	MKA	RHK	MKA	GYN / []
D4	RHK / []	NCH	ONC / []	NCH / []	NCH
			Cluster E		
E1	[]	[]	CAH / []	[]	CAH
E2	[]	[]	[]	[]	[]
E3	VAT	[]	[]	[]	[]
E4	CAH	VAT / []	VAT	VAT	VAT
E5	CAH	CAH	CAH	CAH	CAH
E6	CAH	CAH	CAH	CAH	CAH

Empty rooms are denoted by '[]'. If week A and week B are identical, then only week A is shown. Otherwise, the backslash indicates the decisions for the two different weeks. This template, except for a few minor modifications, is identical to the actual MSS used at the hospital. OR B3, when not booked for TX, will often be used to accommodate non-electives (emergency OR).

estimates and therefore relates to the planning phase. The table also shows that if non-elective demand arriving during high impact periods is deducted from the available OR capacity, then the total slack capacity is reduced to 10%. Interestingly, this is also the value that is suggested to work best by M'Hallah and Al-Roomi [198].

Table 3.6 shows that we allocate to some disciplines an amount of capacity in the model that is different from reality. For instance, the weekly capacities assigned to CAH and VAT were both reduced by 2 ORs (18h). This is done to get a more up to date system as also in reality, from the second half of 2013 on, their assigned capacity decreased.

Table 3.6 Comparison of the slack capacity (hours) used in reality and in the model.

		GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	μ	$\cup$
Reality	open cap.	67.8	46.3	97.5	128	68.0	32.1	55.1	73.8	74.3	57.0	57.1	20.9	46.0	63.4	824
	slack %	8%	87%	21%	19%	8%	22%	31%	7%	6%	9%	47%	14%	3%	22%	20%
	slack % *	6%	58%	3%	11%	-4%	17%	26%	-4%	-5%	4%	39%	10%	0%	14%	10%
Model	open cap.	68.0	18.0	97.5	110	68.1	32.1	55.3	74.0	74.5	57.2	39.3	21.0	46.1	58.5	761
	slack %	9%	67%	21%	6%	8%	22%	32%	7%	6%	9%	23%	15%	4%	18%	14%

The table shows that the total capacity available at all ORs amounts to on average 824 hours a week. Out of this capacity, 20% is slack capacity. Slack capacity is based on the available OR capacity and the estimated caseload (i.e., caseload based on estimated durations). Slack capacity is part of the MSS and is different from the capacity that is used to protect against overtime at some hospitals (usually inserted at the end of the daily schedule). The slack capacity that remains after reducing the available OR capacity with the expected non-elective caseload is denoted by a '*'. The difference in open capacity in reality and in the model for Tx is only theoretic as Tx electives will in reality on average only be scheduled into around 2 ORs a week (18 hours), the rest (87% of the slack) is generally used to accommodate urgent transplant patients. The difference for VAT and CAH reflects recent changes in the capacity allocation schema.

It is also important from a capacity perspective to create a realistic model of the non-elective OR assignment schema. Figure 3.5 shows that the elective load on different clusters is different. A balanced load on clusters is only observable if non-electives are included. This is fair to do as non-electives usually enter ORs that are assigned to the elective discipline itself or to the cluster of the discipline (Fig. 3.2), i.e., they contribute to the clusters' load. Therefore, in the model it is important to allocate non-electives realistically as, e.g., a random assignment would yield a false load on clusters and disciplines. This would lead to a false view on rescheduling and on OR- and patient-related performances.

3.1.5 Rescheduling

There are two major reasons why regular OR time is not always enough to serve all planned electives. Firstly, it frequently happens that surgeries take longer than estimated (Fig. 3.4). Secondly, a non-elective arrival, generally of DT category 1, can demand immediate access to an OR and thus postpone the execution of the OR's elective schedule (Fig. 3.2). In those cases it can become necessary to reschedule elective patients in order to avoid excessive overtime. The main factor why rescheduling is imperative are nurses. Nurses that work longer on one day will be missing on the next day as they will have to recuperate. This would lead to operational problems.

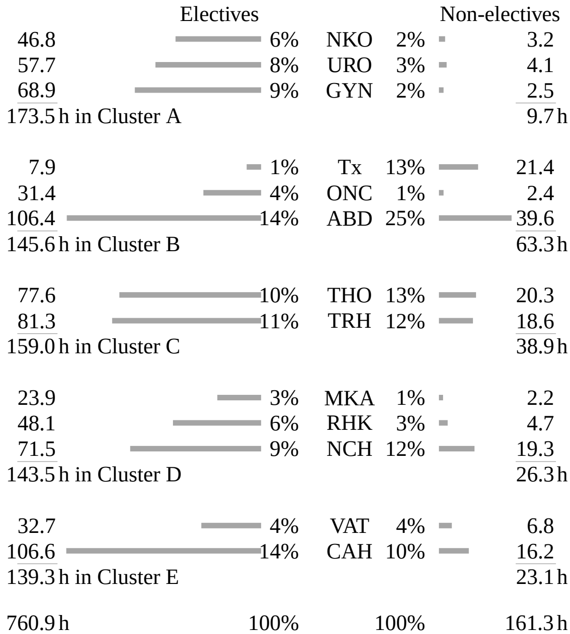


Fig. 3.5 Average weekly capacity used by elective and non-elective patients. The values are based on realized durations. For planning purposes less capacity is booked since duration estimates are generally underestimated (Table 3.6).

We distinguish between two basic types of rescheduling actions: surgery reassignment and surgery cancellation. A surgery is reassigned if, on the day of the surgery, it is moved from the originally planned OR to another OR. The surgery is, however, still performed on the originally planned date. On the contrary, a cancelled surgery will be performed on a later date and is assigned to the surgeon's next OR session. This is done even if the next session is already fully booked.

We make a clear distinction between rescheduling and replanning. Rescheduling is done on the day of the surgery and is used to avoid excessive overtime. It is not part of the patient scheduling process, but a component of the simulation model. We therefore do not test different rescheduling policies and only model the current practice found at the hospital. In contrast, replanning a surgery is done before the surgery date and is therefore part of the scheduling process.

In our setting, we focus on rescheduling as an action to control some aspect of OR overtime. In some hospitals, a limit on the OR overtime is enforced. For instance, in a Spanish setting described by Pulido et al. [228] this limit is two hours of overtime. For work beyond that limit, surgeons and nurses are not paid, giving them an incentive to rather reschedule a surgery than to go over the set time limit. Other hospitals may control the risk of going into overtime as they ensure that an OR goes into overtime only in a certain percent of the cases. This is done, among others, so that nurses will only occasionally have to work longer hours. Other, mostly profit-oriented hospitals may trade off the cost of paying for overtime staffing and the profit gained from performing a surgery in overtime. At the University Hospital Leuven, depending on the hour of the day, a limit is set on the number of ORs that are allowed to stay open in overtime, i.e., 8 ORs out of the 22 ORs may be running overtime beyond 6 pm, 4 beyond 7 pm and only 2 beyond 8 pm. Those 2 ORs remain open the entire night and serve incoming non-elective patients.

At the hospital, cancellations are less often carried out than surgery reassignments. Understandably, the hospital's head anesthesiologist is more reluctant to cancel a surgery than to reassign it to another OR. This is the case as it is frustrating for patients to be cancelled. Patients being reassigned to another OR is not a problem. As a matter of fact, they might not even notice it.

At the hospital, between one and two patients are cancelled on a daily basis and more than six are OR-reassigned. For an elective patient, this means a probability of 3.4% to be cancelled and of 13.1% to be reassigned. The hospital targets a cancellation rate of 2% in the future.

In the literature the term cancellation can also cover actions where surgeries are cleared from the schedule already prior to the surgery day [89]. Using this definition, Dexter et al. [84] show that inpatients are more often cancelled than outpatients. Moreover, Epstein and Dexter [87] conclude that cancellations do not need to be interpreted as a system failure as scheduling cancelled patients does not increase the variability in the OR workload. Therefore, some amount of cancellation is not only unavoidable, but it can also be desirable.

In the simulation model, we imitate the behavior of the hospital's head anesthe-

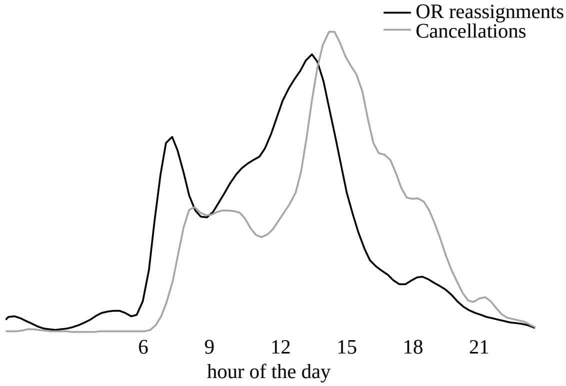


Fig. 3.6 The distribution of the time of day when patients are rescheduled. Rescheduling decisions are made continuously throughout the day. The distribution has two peaks, one in the morning and one in the afternoon. The larger peaks in the afternoon are at 2 pm and 3 pm for OR reassignment and cancellation respectively. Similarly to our findings, also Epstein and Dexter [89] show a histogram with two peaks for cancellations on the day of surgery, one peak in the early morning and one in the afternoon. In their setting, however, a larger peak is present in the morning (around 7 and 8 am) than in the afternoon.

siologist. The head anesthesiologist makes rescheduling decisions continuously throughout the day. Therefore, in the model, each full hour from 8 am to 9 pm we identify those ORs that are expected to run into overtime (Fig. 3.7). From the identified ORs it is then checked whether surgeries can be moved to other ORs. An OR can only accept a surgery if, including the new surgery, the OR is still estimated to close before OR closing time (16.45). In case a surgery cannot be reassigned to another OR it may get cancelled.

In the simulation model, a surgery that is canceled is assigned to the closest date where the same surgeon has an open OR. If there is no such date within one week, also empty ORs can be chosen. A surgery that is canceled once is always served first in the replanned OR, ensuring that it is not canceled again.

Rescheduling actions are based on the estimated closing time of ORs. This estimate is the sum of two components: firstly, the sum of the estimated surgery durations still waiting and, secondly, the amount of time the current surgery is still expected to need. This last component is regarded to be zero in case the surgery already takes longer than its estimate. This is not a correct estimate, but

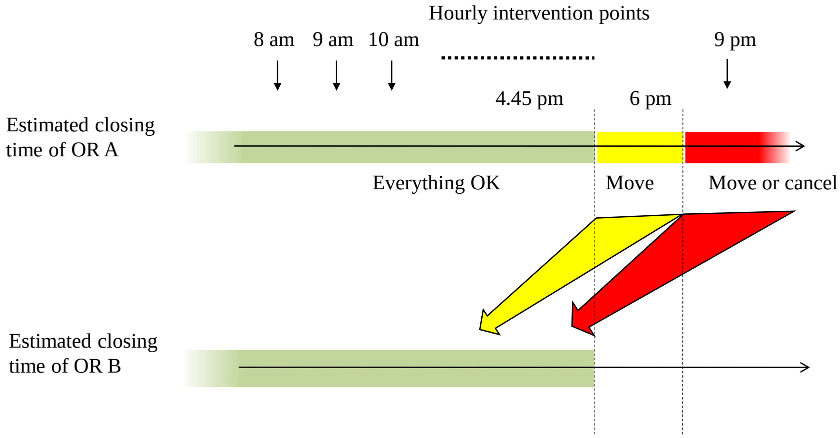


Fig. 3.7 Depending on the estimated closing time of the OR, a surgery can be OR-reassigned or cancelled. The expected closing time of OR A can fall into three intervals: green – OR closes in time, nothing has to be done; Yellow – OR goes overtime, if possible move last surgery, else keep it; Red – OR goes heavily into overtime, if possible move last surgery, else cancel it.

is likely to be close to the value that is used in reality. The correct estimate would be the expected value of the conditional probability distribution of the surgery length given the amount of time the patient is already in the OR.

As shown by Figure 3.6, rescheduling decisions are made continuously from the morning until the evening. Consequently, instead of rescheduling patients at a certain hour of the day, in our model we allow rescheduling interventions to happen on an hourly basis starting from 8 am to 9 pm. At each intervention point, we collect the ORs that are believed to run into overtime. For those ORs, it is attempted to move their last surgery to another less occupied OR.

There is one problem with the method explained in Figure 3.7, namely, that if implemented, a disproportionately large number of surgeries will be rescheduled already early in the morning. This happens in case the first surgery is taking longer than expected.

In reality, whether we believe that an OR goes overtime or not will depend on the degree of trust we put into our estimate. The later we are in the day, the more surgeries have been realized and consequently the better our estimate becomes.

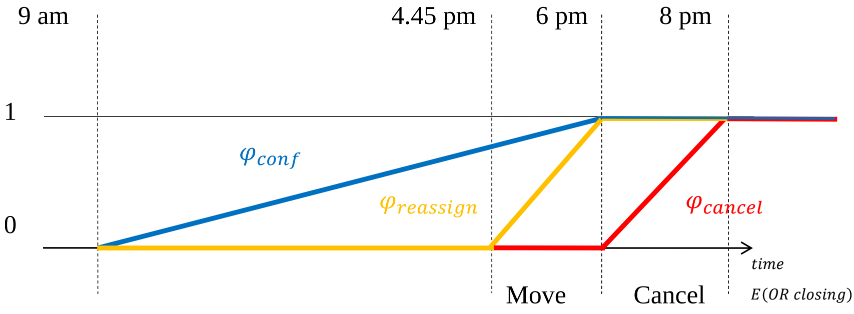
This relationship between the current time and the estimated OR closing time is captured by the formulas in Figure 3.8. In the formulas, the degree of trust we have in our estimates is modeled as a linear function of time (in blue). Put differently, we normalize the time of the day, that is, we map the time of the day onto the unit interval. Similarly we also normalize the estimated OR closing time. Estimates can be normalized in the two ways represented by the functions in yellow and red. The yellow function is used to check whether an OR satisfies the criteria to move its last surgery to another OR, that is, the OR is believed to go into overtime to a degree that justifies an OR reassignment. Similarly, the red function is used to test for cancelation. OR reassignments will never happen if the estimated OR closing time is before 4.45 pm, as the yellow function will be on a constant zero. For the same reason we will never cancel a surgery from an OR that is estimated to close before 6 pm.

Surgeries cannot freely be moved between two ORs. In the simulation model (as in reality) a strict hierarchy is followed (Fig. 3.9). Following this hierarchy allows to reduce the negative impact OR reassignments have on the OR department.

3.2 Model

The algorithms that are most interesting to the hospital are the ones that consider the DT and are manually usable by the surgeons. This is the case as surgeons prefer to plan their patients themselves. This is unlikely to change in the near future as surrendering patient scheduling to a central authority would mean that surgeons would lose part of their independence. At the moment, there is no central hospital-wide patient scheduling system in place.

Additionally, it would be difficult to convince all surgeons of the benefits of using a computer and optimization software to schedule their patients. This is one of the major reasons why formulating the patient scheduling problem as an optimization problem would, in our setting, be of limited use.



OR reassign IF: $\varphi_{conf}(time)* \varphi_{reassign}(E(OR\ closing)) > 0.3$

Cancel IF: $\varphi_{conf}(time)* \varphi_{cancel}(E(OR\ closing)) > 0.5$

Fig. 3.8 The decision if a surgery is OR-reassigned or cancelled depends on a formula that considers the hour of the day and the estimated OR closing time. The last surgery of OR A is reassigned to another OR if, for example, it is 12 noon (blue function takes the value 1/3) and the estimated closing time of OR A is 6 pm (yellow function takes the value 1). The multiplication of 1/3 and 1 results in 1/3 which is larger than the threshold of 0.3. This would not be the case if the intervention point would be checked an hour earlier at 11 am. A similar logic applies to cancellations. The thresholds '0.3' and '0.5' were chosen on a 'trial and error' basis, trying to fit the histograms in Figure 3.6. Noteworthy is the fact that the blue function reaches its maximum at 6 pm which means that after that point in time, we have full confidence in the estimates.

3.2.1 Model assumptions and validation

Our results are usable in the real setting as we ensured that the model is credible and valid. Model credibility is concerned with "developing in potential users the confidence they require in order to use a model and in the information derived from that model" [243]. We created the model based on the data of the University Hospital Leuven and through numerous meetings with the management (head surgeon, head nurse, etc.) it was gradually confirmed that we have the right understanding of both the data and the setting. Our model is consequently credible to the people at the hospital.

Model validation is the "substantiation that a computerized model within its

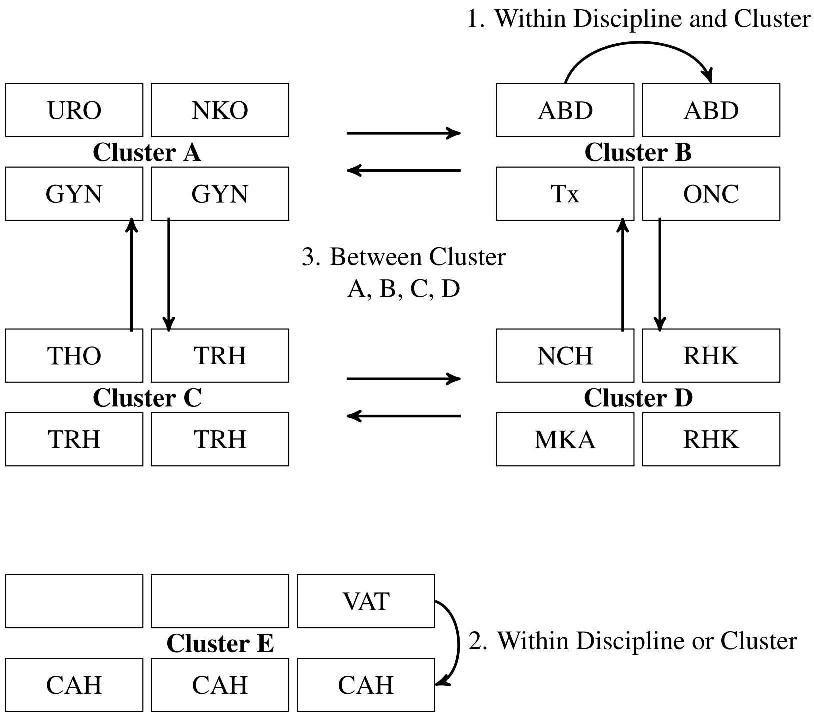


Fig. 3.9 The surgery OR reassignment schema. Firstly, it is preferred to reassign a patient to an OR that serves the same discipline. Alternatively, a surgery can be reassigned to an OR of another discipline as long as it is within the same cluster. Less favorable but possible is to move surgeries across clusters A, B, C and D (Fig. 3.5). CAH can only be reassigned to its own ORs and NCH surgeries cannot leave the cluster to which they are assigned.

domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model" [244]. We validated our model by comparing the simulation results with real hospital-data. We think that there are three aspects that are of key importance and thus have to be validated: (1) to realistically allocate capacity to disciplines (Table 3.6), (2) to realistically model the arrival caseload (Table 3.7) and (3) the validation of key hospital-related performance measures (Fig. 3.10). As we already dealt with the first point in Section 3.1, we will only focus in this section on points 2 and 3.

The arrival caseload is the amount of OR hours that have to be scheduled for surgery in the current or future weeks, which is different from the planned surgery caseload assigned to those weeks. Whereas the surgery caseload depends largely on the fixed MSS and is thus fairly stable and predictable, the arrival caseload is more variable as it depends both on stochastic patient arrival numbers and on their stochastic surgery duration lengths.

Table 3.7 shows that the model is realistic as both averages and standard deviations reflect reality (small Δ). This is true for the arrival caseload based on realized durations, for the caseload based on estimated durations and for the error between them. The error is important as it contributes to the uncertainty that differentiates a planned from a realized schedule.

In order to ensure that the hospital processes are modeled accurately, we validate some of the key hospital-related performance measures (Fig. 3.10). The results confirm that the model is valid. There are four measures where there is a statistical difference. In reality (1) a lower utilization, (2) less overtime, (3) more undertime and (4) more daily open ORs are experienced.

In case of measures (2) and (3), the difference is statistically speaking significant, but is small enough to not be of practical relevance for the hospital. The reasons for (1) and (4) can be explained by the fact that in reality less than 9 hours of capacity might be allocated to an OR-day. However, as it is difficult to identify those OR-days, we will assume they are always assigned the full 9 hours. The measured utilization in reality will thus be lower than in the model. Similarly, as in the model and in reality the same amount of total capacity is used, shorter opening hours entail that the OR is open on more days.

3.2.2 Simplifications

We tried to simplify the arrival model, without success. For instance, modeling patient arrivals with a Poisson distribution (as often done in the literature) causes the standard deviation of the average caseload per discipline to decrease from 20.2 hours (measured in reality) to 15.7 hours. For some disciplines, it will be 8 hours lower than in reality, which is almost an entire OR. This means that using a Poisson arrival process may lead to a system that is much more stable than

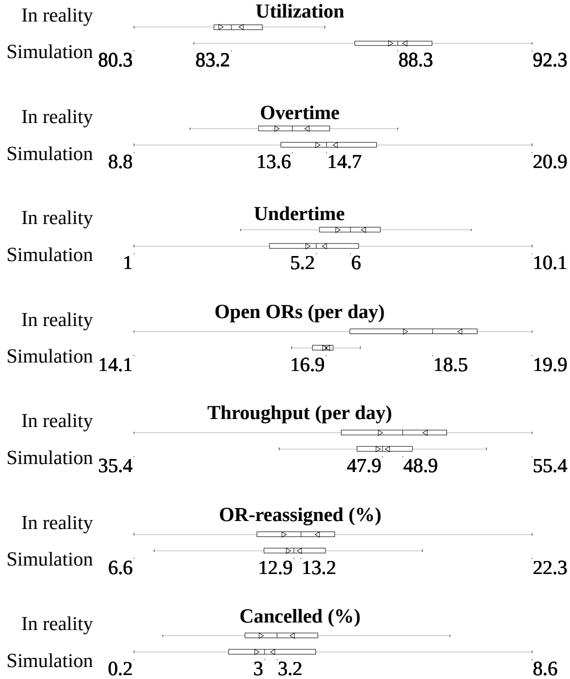


Fig. 3.10 The results of the validation of the simulation model. The central mark represents the median and the edges of the boxes the 25th and the 75th percentiles. The triangular markers approximate the 95% confidence interval. If these intervals do not overlap, then we regard the two medians to be significantly different at the 5% significance level.

it is in reality. This can lead to misleading results. This is especially true with regards to DT 4 patients which, as they are required to be served within a week, are more sensitive to short-term capacity shortages.

Other simplifications involve the surgery duration model, where we tried to model durations in the univariate space and fitted a parametric distribution on both realized and estimated durations independently. The chosen parametric distribution is, for each discipline, the one with the best AIC value. The tested distributions include, amongst others, the ones described in Table 3.3. Interestingly, this would lead to good results with regards to the average estimation error of the weekly caseload. The problem however is the standard deviation, which increases to 14.6 hours from the 5.5 hours measured in reality.

We also tested whether a model using a univariate distribution for the realized durations and a univariate distribution for the error between realized an estimated durations would bring the desired result (estimated durations are then the sum of the two). On the positive side, this method generally gives smaller errors than if any of the other previously mentioned simplifications are used. On the negative side, there will be extreme cases as, for instance, the estimated weekly arrival caseload for RHK would show a standard deviation of 42.2 hours instead of the 23.2 hours measured in reality. This is especially a problem for RHK as a large part of their patient population belongs to DT 4. We consequently think that it is not possible to include any of the previously mentioned simplifications without substantially changing the setting.

3.2.3 Details on the used DES model

We found that only a DES model is able to realistically capture all the aspects of the University Hospital Leuven's scheduling setting that we deemed to be important. The DES model incorporates all the aspects of the surgery setting of the University Hospital Leuven that we found to be vital. We included aspects that relate to the way surgeries are scheduled and replanned before the surgery date and are rescheduled on the surgery day itself. We also replicated the functions of the OR department. This includes, for instance, an implementation of the non-elective to OR allocation schema.

DES models are often created either using custom healthcare-related simulation software or from scratch using a general purpose language. Both methods have their benefits, but they also have their drawbacks. The first option is quick to implement, but there are mechanisms that are difficult to model within the software. Additionally, it is usually a standalone software and thus it is tedious and often impossible to properly integrate with other environments. The second option, using a general purpose language, has the drawback of being time consuming to implement. Advantages are: flexibility, speed and high integrateability.

We chose to create a simulation model that is based on a general purpose language, but is seamlessly integrated with a custom DES environment. We are using Matlab for routines and Simulink's SimEvents toolbox for the DES framework.

An advantage of using Matlab together with Simulink is that components from each environment are easily integrated. Practically, this means that we are able to call the Simulink DES model from Matlab while within Simulink we are able to use Matlab code.

SimEvents is used to simulate the patient service process. This involves invoking appropriate scheduling methods for elective patients. Also certain hospital mechanisms are implemented in SimEvents such as the surgery process in the OR and patient rescheduling actions. One of the drawbacks of SimEvents, in comparison to other DES environments, is its rudimentary nature, making it difficult to directly implement more complex mechanisms. This is to some extent compensated for by the fact that different models from other Simulink environments (e.g., state machines) or Matlab code can be mixed into SimEvents. One of the strong sides of SimEvents is that an entity (patient) can enter an attribute function block. The strength of this block comes from the fact that within the block it is, on the one hand, possible to arbitrarily change the entities' attributes (e.g., scheduled OR) and, on the other hand, it allows to import object handles. This means that patient attributes can freely be processed within the simulation model using any class function created in Matlab. This practically means that an entity can be processed in an arbitrary way, resulting in a highly flexible and capable environment.

We analyzed and imitated the real mechanisms encountered at the hospital. We made a minimal amount of modeling assumptions and used real data as the basis of all submodels. In cases where the data did not reveal enough about a process, we were helped by our contacts at the hospital who provided us with the missing knowledge.

The attributes of patients generated in the model realistically reflect the attributes of the inpatient population of the hospital. Patient attributes are: surgeon ID, arrival rate for each weekday, estimated and realized surgery duration and DT category. Discipline-related attributes are: surgery start time bias (for the first surgeries of the day) and turnaround time.

The statistics for patient attributes are measured for each discipline separately. In the model, all patient attributes are generated on the basis of empirical distributions. Exceptions are the non-elective inter-arrival times and the realized and estimated surgery durations. Non-elective inter-arrival times, for modeling purposes, are assumed to follow the exponential distribution for a given period. A period depends on the weekday and the daytime (daytime: 6 am to 10 pm, nighttime: 10 pm to 6 am). The relation between the estimated and the realized surgery durations is modeled using a statistic that is based on copulas [274] as described in Section 3.1.3.

In order to be aligned with the two-week MSS cycle used at the hospital, all performance measures are batched on a two-week period basis, i.e., one batch covers two-weeks. Mean values shown in the results will therefore be the mean values of individual two-week batch means. Similarly, also the shown standard deviations will relate to the variability between those batches. The simulation length is set to 4000 days.

Batches are only formed from a simulation interval where the measured surgery throughput in all simulation scenarios is within 5% compared to the mean of all scenarios (i.e., warm-up and cool-down periods are removed). Determining the start and end point of this simulation interval can be strongly influenced by the daily fluctuation of the throughput observed within scenarios. To compensate for these fluctuations we smooth out the throughput and apply a low-pass filter with coefficients equal to 1/19 (i.e., we take the moving average with a span width equal to 19 days).

From the 4000 simulation days the valid simulation interval is formed by removing around 30 days from the beginning and around 120 days from the ending. The remaining days form the approximately 275 batches used in the result section.

Using speeding up procedures such as look-up tables and state-flow machines for the OR logic, the run time for each scenario is kept under 1.5 hours on an Intel Xeon, x5690 (3.47Ghz, 64 GB memory). As scenarios can be run in parallel, computation times do not pose a big problem.

3.3 Conclusion

In this chapter we discussed the OR scheduling setting of the University Hospital Leuven and described how we modeled it. We have shown that the weekly and daily arrival variability is high for most disciplines (Table 3.1). We have also shown that non-electives are not allocated to a randomly chosen OR but are instead served in an OR that is assigned to the discipline of the non-elective patient (Fig. 3.2).

We found that for surgery durations modeled on the pathology level, the log-logistic distribution should be used (Table 3.3). If realized durations are modeled on an aggregated level, we advice to use a GMM. If both realized and estimated durations are modeled, then either a fully empirical distribution should be chosen or a bivariate copula model that is able to handle multimodality (Table 3.4).

We introduced the predetermined and static MSS and showed how elective rescheduling is included into the simulation model. Elective rescheduling contributes to the fact that overtime will not depend on the chosen patient scheduling strategy contrary to what is often assumed in the literature. Finally, we discussed details on the simulation model and concluded that computational times do not pose a big problem.

Table 3.7 The caseload of weekly arrivals in reality and in the model.

		Non-Elective	Elective														
			GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	μ	$\cup$
Realized (h)																	
real	μ	**160**	**68.6**	**7.8**	**106**	**106**	**71.2**	**31.3**	**48.0**	**77.3**	**80.9**	**57.5**	**32.6**	**23.8**	**46.6**	**58.3**	**758**
	σ	*31.5*	*20.7*	*6.0*	*28.3*	*39.1*	*27.8*	*11.8*	*23.2*	*28.5*	*19.3*	*15.2*	*14.3*	*14.6*	*14.1*	*20.2*	*138*
	CV	.20	.30	.76	.27	.37	.39	.38	.48	.37	.24	.26	.44	.61	.30	.40	.18
model	μ	**159**	**69.7**	**7.6**	**106**	**105**	**71.1**	**31.8**	**49.0**	**78.2**	**80.5**	**57.6**	**32.7**	**23.9**	**46.5**	**58.5**	**760**
	σ	*26.4*	*21.5*	*5.5*	*28.3*	*36.6*	*28.2*	*11.6*	*23.6*	*27.4*	*19.3*	*15.1*	*14.2*	*15.1*	*15.1*	*20.1*	*80.3*
	CV	.17	.31	.73	.27	.35	.40	.36	.48	.35	.24	.26	.43	.63	.33	.40	.11
Δ	μ	**.97**	**-1.0**	**.28**	**-.09**	**.79**	**.10**	**-.47**	**-1.1**	**-.89**	**.36**	**-.06**	**-.13**	**-.14**	**.14**	**-.17**	**-2.2**
	σ	*5.2*	*-.80*	*.44*	*.02*	*2.5*	*-.39*	*.25*	*-.36*	*1.0*	*.03*	*.02*	*.13*	*-.46*	*-1.0*	*.11*	*57.8*
	CV	.03	-.01	.03	.00	.02	-.01	.01	.00	.02	.00	.00	.01	-.02	-.02	.00	.08
Est. (h)																	
real	μ	**141**	**62.1**	**5.9**	**77.3**	**103**	**62.8**	**25.1**	**37.9**	**68.5**	**70.1**	**52.0**	**30.3**	**18.0**	**44.4**	**50.6**	**657**
	σ	*26.3*	*18.7*	*4.1*	*21.4*	*37.1*	*23.9*	*9.9*	*18.8*	*25.3*	*16.2*	*13.8*	*13.2*	*10.8*	*13.0*	*17.4*	*113*
	CV	.19	.30	.70	.28	.36	.38	.40	.50	.37	.23	.27	.44	.60	.29	.39	.17
model	μ	**138**	**62.5**	**5.7**	**77.2**	**102**	**62.4**	**25.1**	**38.4**	**68.8**	**69.3**	**51.8**	**30.1**	**18.0**	**44.1**	**50.4**	**656**
	σ	*23.2*	*19.5*	*4.1*	*20.4*	*35.4*	*24.3*	*9.0*	*18.9*	*23.8*	*16.0*	*13.2*	*12.9*	*10.9*	*13.7*	*17.1*	*69.3*
	CV	.17	.31	.72	.26	.35	.39	.36	.49	.35	.23	.26	.43	.60	.31	.39	.11
Δ	μ	**2.1**	**-.39**	**.20**	**.04**	**.90**	**.40**	**-.06**	**-.59**	**-.28**	**.74**	**.21**	**.13**	**-.09**	**.34**	**.12**	**1.5**
	σ	*3.1*	*-.84*	*.05*	*.96*	*1.7*	*-.33*	*.88*	*-.15*	*1.5*	*.19*	*.58*	*.31*	*-.07*	*-.78*	*.31*	*44.2*
	CV	.02	-.01	-.02	.01	.01	-.01	.04	.00	.02	.00	.01	.01	.00	-.02	.00	.07
Est. error (h)																	
real	μ	**19.7**	**6.5**	**1.9**	**28.7**	**3.1**	**8.4**	**6.3**	**10.1**	**8.8**	**10.9**	**5.5**	**2.3**	**5.8**	**2.3**	**7.7**	**101**
	σ	*11.5*	*6.2*	*2.5*	*10.1*	*5.8*	*6.6*	*4.2*	*6.1*	*6.3*	*6.7*	*4.0*	*4.0*	*5.0*	*4.4*	*5.5*	*32.7*
	CV	.58	.96	1.3	.35	1.9	.78	.67	.60	.71	.62	.73	1.7	.86	2.0	1.0	.33
model	μ	**20.9**	**7.2**	**1.9**	**28.8**	**3.2**	**8.7**	**6.7**	**10.6**	**9.4**	**11.2**	**5.8**	**2.6**	**5.9**	**2.5**	**8.0**	**104**
	σ	*9.3*	*5.6*	*2.4*	*10.3*	*5.5*	*6.9*	*4.5*	*6.2*	*6.7*	*6.5*	*4.1*	*4.7*	*5.6*	*4.6*	*5.7*	*22.4*
	CV	.45	.79	1.3	.36	1.7	.79	.67	.59	.71	.58	.72	1.8	.95	1.9	.99	.22
Δ	μ	**-1.2**	**-.64**	**.08**	**-.13**	**-.11**	**-.30**	**-.41**	**-.48**	**-.61**	**-.38**	**-.27**	**-.27**	**-.04**	**-.20**	**-.29**	**-3.7**
	σ	*2.2*	*.60*	*.16*	*-.26*	*.32*	*-.32*	*-.26*	*-.16*	*-.42*	*.25*	*-.14*	*-.69*	*-.59*	*-.15*	*-.13*	*10.2*
	CV	.14	.17	.03	-.01	.17	-.01	.00	.01	.00	.04	.01	-.09	-.09	.10	.03	.11

The difference between the modeled arrival caseload and the real arrival caseload is small for non-electives, elective disciplines and the combination of all patients.

Chapter 4

One-step strategy

In our model, we imitate the reality of the hospital where patients are scheduled to a final surgery date during their consultation session. The surgeon or the administrative people with the input from the surgeon find a suitable date and OR without a scheduling algorithm. Only those dates are considered on which the surgeon is assigned an OR. In the simulation model, as in reality, we ensure that, firstly, only patients associated to the same surgeon can be assigned to a particular OR and, secondly, a surgeon can only be assigned to one OR a day.

At the University Hospital Leuven an OR can be entirely filled up but is preferably not overbooked. However, there will be disciplines that occasionally overbook for a few hours. This is particularly true for CAH, NCH, THO, TRH, URO and NKO. In the model these disciplines are allowed to overbook, CAH by 2 hours and the remaining five by 1 hour. All other disciplines cannot overbook, i.e., the sum of the expected surgery durations assigned to their ORs cannot exceed 9 hours.

Booking rules can vary from hospital to hospital. At some hospitals, ORs may never be fully booked or, conversely, can be overbooked. For example, at the Erasmus Medical Center in the Netherlands ORs are not fully booked and slack time is scheduled. This ensures that the probability of overtime stays below a certain level [126].

In our setting, a surgery schedule is not necessarily fixed as surgeries can be replanned before the day of their surgery. Surgery replanning to earlier surgery dates can, for instance, be used to improve the usage of ORs. In reality, this is applied to 5.2% of the total patient population. Other reasons why surgeries are brought forward in the schedule are the worsening health condition of the patient or hospital-related logistical reasons. We will focus on OR usage related advantages and investigate whether patient replanning can help utilizing unclaimed free short-term OR capacity. As the hospital generally tries to avoid excessive replanning, we also investigate whether the unused OR capacity can be filled up with new arrivals.

4.1 Factors

We will refer to different scheduling policies or methods as scheduling factors. The combination of those factors creates scenarios, which we then test in the DES model described in Section 3.2.3.

We grouped the different aspects of the patient scheduling process into three factors (Table 4.1). The first factor tests the use of the first come, first served (FCFS) strategy which assigns patients to the earliest possible surgery date regardless of their actual DT. The second factor tests the use of pushing lower urgency patients into the future leaving capacity free for higher urgency DT categories. The third factor tests the use of filling up unclaimed short-term free capacity. This is tested in two ways, firstly, by using patients arriving one day in advance, i.e., patients are allowed to be scheduled to exactly one day after their arrival, and secondly, by replanning patients from future dates to earlier dates. Replanning is done before any new elective arrival is registered for the current day.

Table 4.1 Tested one-step scheduling factors.

	Factor	Values
F1	FCFS	None, $\leq$ DT 4/5/6/7/8
F2	DT interval	Early, center, late
F3	Next day	None
	Fresh arrivals	$\leq$ DT 4/5/6/7/8
	Replanning patients	$\leq$ DT 4/5/6/7/8, APQ

Each factor can take several values. The combination of the three factors forms a scheduling scenario. For example, a scheduling scenario is to serve patients up to DT category 5 on a FCFS basis, schedule the rest (DT categories 6 to 8) to the center of their respective DT interval and fill up next day capacity with the APQ.

4.1.1 Factor 1: First come, first served

Factor 1 (F1) is used to investigate whether it is beneficial to allow patients, up to certain DT categories, to be served FCFS. As Table 4.1 shows, the factor can take 6 values: (1) none of the patients are served on a FCFS basis, (2) only DT category 4, (3) DT categories 4 and 5, (4) DT categories 4 to 6, (5) DT categories 4 to 7 or (6) DT categories 4 to 8 patients are served FCFS. The factor allows patients of the included DT categories to be served as quickly as possible. Indirectly, it also means that the patients of all DT categories served FCFS are treated equally. For example, if FCFS applies to patients up to DT category 6, then, from a scheduling perspective, DT category 5 and 6 are regarded to be equally urgent as DT category 4.

As also proposed by Vijayakumar et al. [293] and Niu et al. [207] patients served FCFS will be assigned to the first date that has a suitable open OR available. In case that such an OR is not available, a new OR is opened. In our model, surgeries can only be allocated to ORs that are assigned to the corresponding discipline and to the corresponding surgeon. A surgery can be allocated to a new empty OR if the OR is assigned to the respective discipline in the MSS. The newly opened OR will be assigned to the surgery's surgeon and only accept those additional future surgeries that belong to the same surgeon.

4.1.2 Factor 2: DT interval

Factor 2 (F2) is used to postpone less urgent surgeries, thereby creating short-term buffer capacity that can be used by more urgent patients. There are three strategies: schedule patients into the early, center or late part of their DT interval. With the early strategy, patients are assigned into the closer beginning of their DT interval. This is similar to FCFS with the restriction that patients can only be served after the start of their DT interval. With the center strategy, patients are scheduled as close to the middle of their DT interval as possible, i.e., the temporal distance between the selected date and (DT end + DT start) / 2 is minimized. With the late strategy, patients are scheduled into the end of their DT interval, that is, patients are served as late as possible within their DT. If there is no such date available, then a date after the patient's DT is chosen.

It is interesting to explicitly incorporate the DT into a scheduling strategy as serving patients closer to their due date is a concept that can intuitively feel advantageous to surgeons. This approach is also tested by Rizk and Arnaout [234].

4.1.3 Factor 3: Next day

Factor 3 (F3) is used to quantify the benefits of filling up unclaimed free short-term capacity. This is capacity that in the morning of the preceding day is still shown to be unclaimed and is therefore regarded to be in danger of being wasted. For example, if Wednesday morning the OR plan for Thursday shows 5 hours of unclaimed capacity, then 5 hours of OR capacity are in danger of being wasted. We will refer to this type of capacity as next day free capacity. Next day free capacity can be occupied by patients from two different sources: firstly, new arrivals and, secondly, replanned patients. Replanning works similarly to a waiting list, where the replanning policy determines which patients to pick first from the list [191]. Next day free capacity in an OR is only made available to surgeries that are assigned to the surgeon of that OR.

Factor 3, similarly to factor 1, applies to certain DT categories. It is applied to each discipline separately. In case of replanning, it is used in combination with the best-fit strategy. This means that from the list of eligible patients assigned to future dates (waiting list), those patients are replanned that make best use of the available free capacity. We use a replanning routine that is likely to be most often used in reality. We first replan the patient with the longest estimated surgery that still fits the next day free capacity. The second patient chosen will need to fit the remaining capacity. We continue this process until the left over free capacity does not allow to accommodate any further patient.

Next to best-fit, we also implemented a patient selection strategy that is based on an accumulating priority queue (APQ). In the APQ, patients accumulate priority as a linear function of their time in the queue and their priority [153, 261], i.e., their waiting time and their DT. The weight v_i associated to each patient is therefore:

$$v_i = \frac{(s_i - a_i)}{dt_i} \tag{4.1}$$

where a_i is the arrival day, s_i is the surgery day and dt_i is the DT in days of patient $i \in I$.

4.2 Results

Results are easiest analyzed if the effect of each factor representing a scheduling method can be isolated from the other two factors (i.e., averaged out). This corresponds to an analysis of the main individual effects. Such an analysis can yield misleading results in the presence of interaction effects between any two factors and the particular performance measure. Unfortunately, we noticed that with most of these performance measures there are significant two-way (and three-way) interaction effects present at a 5% significance level. Therefore, we analyzed factors in combination. This can be done with interaction plots (e.g., Fig. 4.1). In order to represent all results in a consistent way, we use interaction plots for all results (also for performance measures where no significant interaction effect was present). As the interaction plots show, a full factorial design with 216 scenarios is used: 6 (factor 1) * 3 (factor 2) * 12 (factor 3).

4.2.1 OR-related performance measures

We distinguish between OR-related performance measures (e.g., utilization, overtime and undertime and patient-related performance measures (e.g., patient waiting time and the ratio of patients that are served within their DT).

Results for each performance measures are visualized using a three-way interaction plot. Each interaction plot contains four dimensions, three correspond to the three scheduling factors (Table 4.1) and one to the respective performance value (y axis). Each point represents one scenario, thus a combination of the three factors.

Table 4.2 shows that many of the scheduling factors have a significant main effect on OR-related performance measures (the p-values are smaller than 0.05). Nevertheless, as the standard deviation between different scenarios is very small, those effects do not bear any practical relevance. A similar observation can be made in Figure 4.1 for overtime. The figure shows that overtime values change between 14.6% and 15.1%. Small differences like these are for the hospital of little practical importance.

The reason why overtime does not depend on the chosen scheduling strategy can be explained as follows. As demand closely matches supply and as surgeons can fill up their ORs fully, ORs will irrespective of the patient scheduling strategy be fully booked and therefore highly utilized (87.9-88.1%). The fact that the ORs are highly utilized is bound to lead to a substantial amount of average overtime. The exact amount is however independent of the tested patient scheduling strategy (Fig. 4.1) but determined by other factors such as the estimation error of surgery durations (Table 3.2). This also means that overtime might not be avoidable without having to sacrifice the efficient use of OR time. And vice versa, OR time might not be efficiently used without overtime.

Since scheduling factors practically speaking do not affect OR-related performance measures, they are excluded from any further analysis. This allows us to concentrate on patient-related performance measures only (Table 4.3). We will focus on three in particular (Fig. 4.2): the percentage of patients served within their DT, the average patient waiting time and the weighted DT cost.

Table 4.2 OR-related performance measures.

		Utiliz-ation %	Over time %	Under-time %	Open OR Daily mean	Throughput Daily mean	Reassign-ed %	Cancel-led %
Real	μ	83.4	13.6	6.0	18.0	47.9	13.1	3.4
	σ	1.3	1.6	1.2	1.5	4.2	2.8	1.1
	CV	.02	.11	.19	.08	.09	.21	.33
FCFS	μ	88.0	14.8	5.2	16.9	48.0	12.9	3.2
	σ	1.8	2.1	1.5	.17	2.0	1.8	1.4
	CV	.02	.14	.29	.01	.04	.14	.45
Scen. avg.	μ	88.0	14.8	5.3	16.9	48.0	12.9	3.2
	σ	.05	.07	.10	.01	.03	.17	.08
	CV	.00	.00	.02	.00	.00	.01	.02
Scen. std.	μ	1.65	1.97	1.47	0.17	1.83	1.67	1.44
Main effect	F1	.96	<.001	<.001	<.001	1.0	<.001	.02
(p-value)	F2	.01	.06	.00	.01	.40	<.001	<.001
	F3	.00	.54	.03	.71	.19	<.001	.21

Real values are compared, in specific, to the FCFS strategy and, in general, to all scenarios. The standard deviation of the scenario means (e.g., 0.05 for utilization) has a different meaning than the mean of the scenario standard deviation (e.g., 1.65 for utilization). The former shows how much the value of a performance measure differs between different scenarios, whereas the latter shows how much variability is generally present between the batches within scenarios. The means in the FCFS strategy are identical or very close to the means of the scenario average. As the standard deviations of the scenario averages are low, the mean values in most other strategies will also be close to the FCFS strategy.

4.2.2 Percentage of patients served within their DT

An important indicator for the hospital is the percentage of patients who are served within the medically advised time limits (the DT) set by their surgeons. Whether patients are served within their DT depends on their arrival and surgery date. For example, if it is determined on a Monday that a DT category 4 patient needs surgery, then the latest date that is within the DT is the Monday of the following week. Later days, regardless of the exact number of days late, are considered to be after the DT. As DT category 8 patients are not given any hard deadline, they are excluded from the calculations.

Currently at the hospital, around 65% of the patients are served within their DT. Further decomposing this result by DT category shows that 81.2% of DT category 4 patients are served within their DT, making it the most efficiently

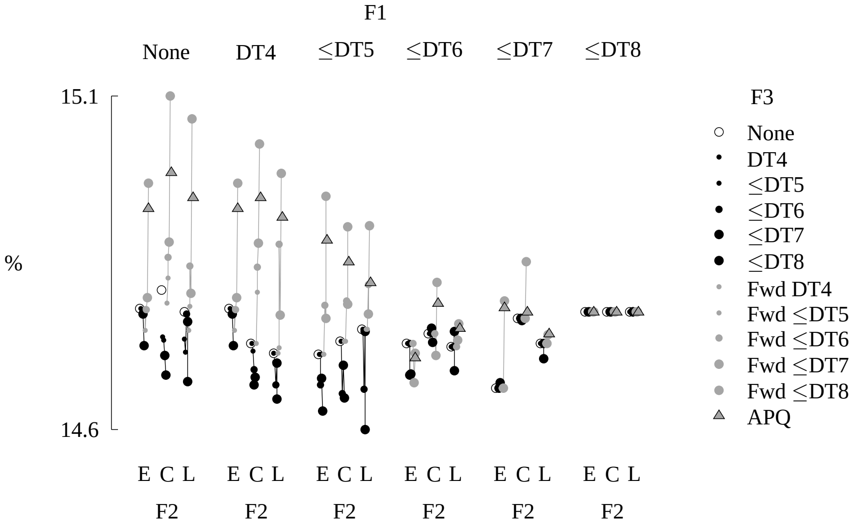


Fig. 4.1 The amount of overtime is, from a practical perspective, independent of the chosen patient scheduling strategy as the minimum and maximum values are very close. Each point in the figure represents a scenario, that is, a combination of three factors. The exact realization of each factor 1-3 is defined by: the label on the top (F1), the label on the bottom (F2) and the marker (F3). The full markers represent scenarios where fresh arrivals (black) or existing patients (gray) are used for replanning. For example, the most left triangle represents a scenario where: no DT category is served FCFS (F1), surgeries are served in the early part of their DT interval (F2) and replanning uses the APQ method (F3). The y axis shows the respective performance value, which in this case is overtime.

served DT category at the hospital. For patients with DT categories 5, 6 and 7, the respective values are 52.6%, 59.5% and 67.2%.

From Figure 4.2 we see that as more DT categories are served FCFS (Factor 1), the percentage of patients served within their DT is increasing. This means that serving patients from a specific DT category FCFS is beneficial as, on the one hand, it naturally decreases access times for patients from the specific DT category and, on the other hand, seems to have at most a limited detrimental effect on patients from other DT categories.

Similarly, also scheduling patients 'next day' (Factor 3) is beneficial. This implies that it is crucial to save capacity that might remain unused. OR capacity

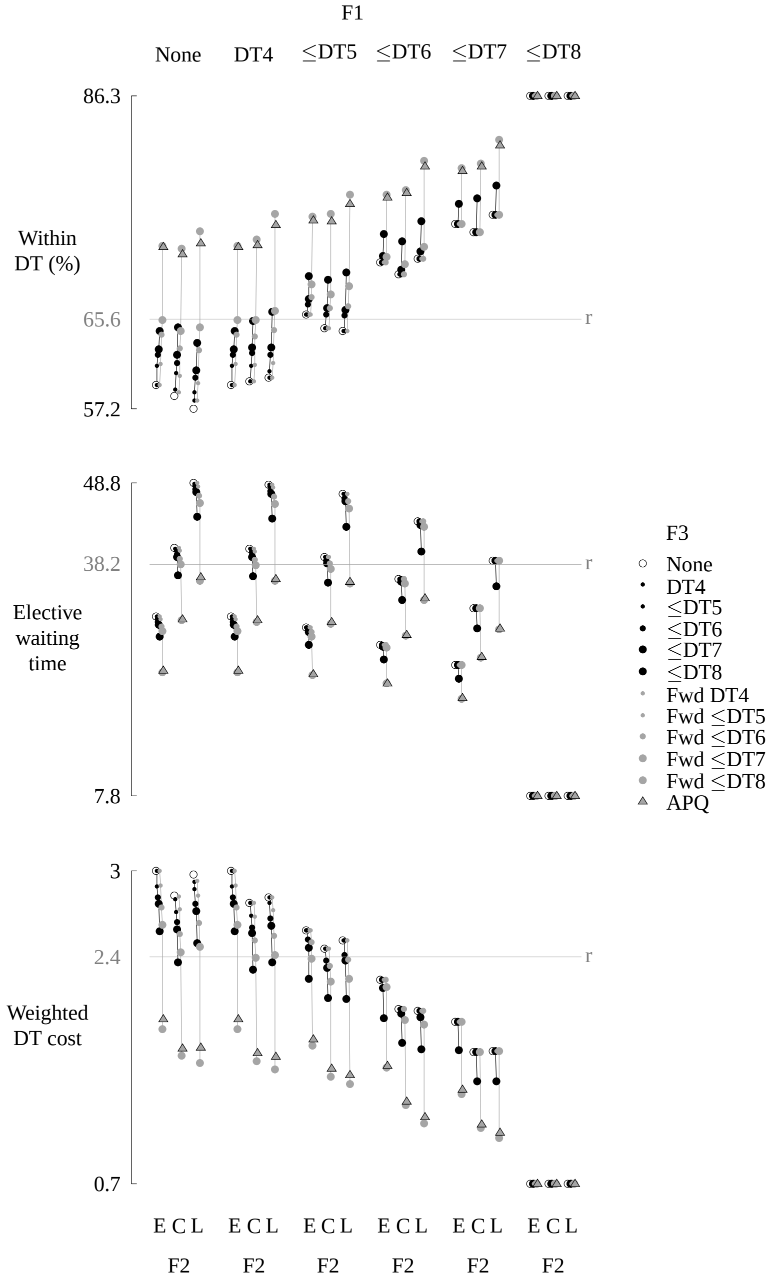


Fig. 4.2 Patient-related performance measures. The horizontal line shows the real values measured at the hospital ('r'). The weighted DT cost is defined in Section 4.2.4.

Table 4.3 Patient-related performance measures.

		DT offset (%)	Elective waiting time	Weighted DT cost
Real	μ	65.6	38.2	2.4
	σ	4.4	4.0	.43
	CV	.07	.11	.18
FCFS	μ	86.3	7.8	.71
	σ	5.3	1.1	.32
	CV	.06	.14	.46
Scen. avg.	μ	71.5	30.8	1.9
	σ	8.7	12.4	.71
	CV	.12	.40	.37
Scen. std.	μ	5.97	1.86	0.55
Main effect (p-value)	F1	<.001	<.001	<.001
	F2	<.001	<.001	<.001
	F3	<.001	<.001	<.001

All factors have a significant main effect on all three performance measures. Additionally, the scenario means show a large CV, which means that factors also practically speaking have a large influence on the results. The fact that the average standard deviation within scenarios (indicated by 'Scenario std.') is high for all three measures, shows that there can be large differences between different two-week periods (see Sec. 4.2.4).

that remains unused is wasted and cannot be recovered anymore, scheduling patients in the very last moment into this capacity avoids that the replanned patients occupy future OR capacity that might be needed for other patients.

Figure 4.2 also shows that filling up next day free capacity by replanning and thus bringing patients forward in the schedule is considerably more effective than using fresh arrivals as an OR can only serve surgeries from one surgeon. This restricts the number of usable new arrivals to a limited set unless one purposely schedules consultations for a surgeon on the day before the day an OR is assigned to that surgeon. More patients are available for replanning.

In contrast, factor 2 shows only a minimal effect. This means that scheduling patients into either the early, center or late part of their DT will result in a similar performance value.

4.2.3 Patient waiting time

The patient waiting time is one of the classical performance metrics used in the literature. The waiting time of a patient equals the number of days between the date the decision for surgery was made and the date the surgery was performed. The decision for surgery is made when the surgeon and the patient meet for consultation and a form is filled out with the details of the surgery. This type of waiting time, usually measured in days, is also called indirect waiting time.

Elective patients at the University Hospital Leuven wait 38.2 days on average (Fig. 4.2). A further decomposition by DT category shows that for DT categories 4, 5, 6, 7 and 8 it is 8 days, 21.6 days, 40.1 days, 51.7 days and 75.1 days respectively (Table 4.4).

Improvements with regards to waiting time can be achieved in an intuitive and straightforward way. Firstly, by scheduling more DT categories on a FCFS basis. Secondly, by serving patients in the earlier part of their DT. Thirdly, by allowing patients to be served next day. Replanning patients is also with regards to the waiting time more effective than using fresh arrivals. As shown by Figure 4.3, lower waiting times apply to all DT categories.

From Figure 4.2 it is interesting to note that improvements generally remain severely limited as long as FCFS (or next day) is only applied to patients up to DT category 7. Real benefits are only realized once DT category 8 patients are included. This shows that the way DT category 8 patients are scheduled is important, but it also highlights some of the drawbacks of using patient waiting time as a performance measure as it can be heavily determined by low urgency patients.

4.2.4 Weighted DT cost

The degree to which the DT is obeyed can be measured in several ways. A straightforward method is to simply determine the ratio of patients who are served within their DT. However, this does not provide any information on the

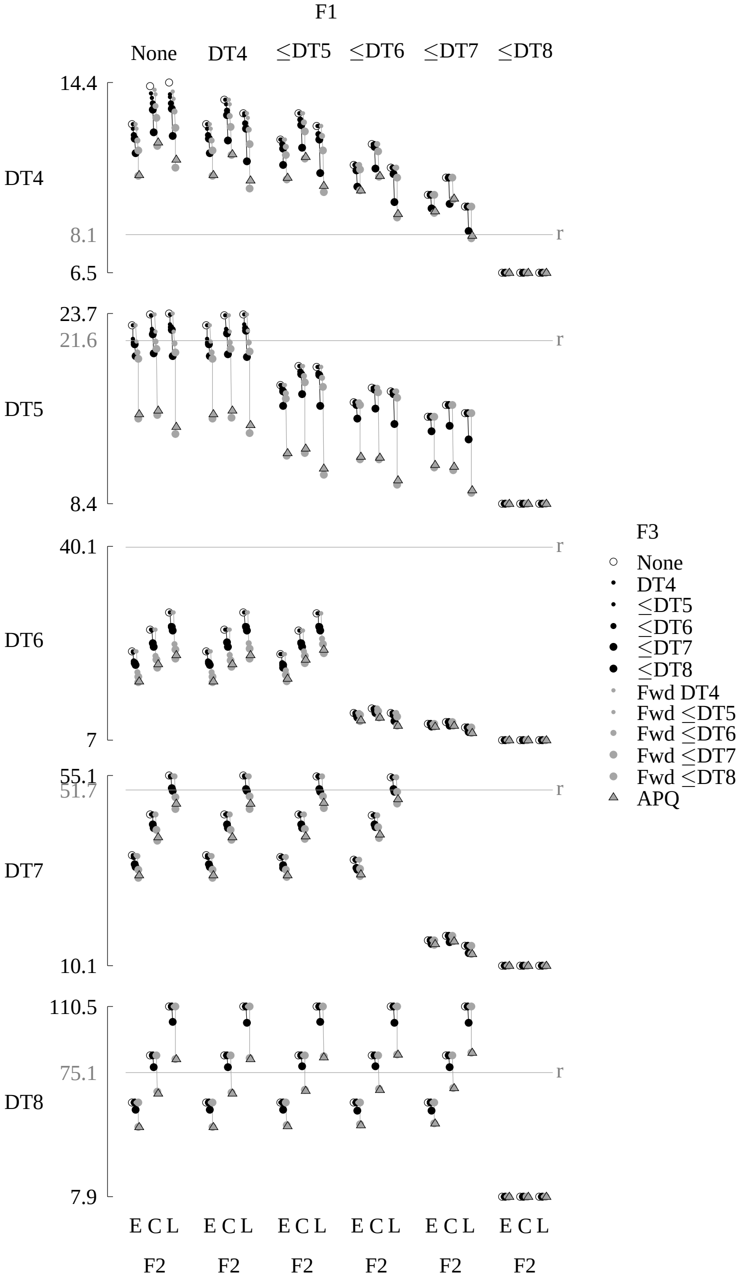


Fig. 4.3 The waiting time of each elective DT category.

Table 4.4 Waiting time by DT category.

		DT4	DT5	DT6	DT7	DT8
Real	μ	8.1	21.6	40.1	51.7	75.1
	σ	1.8	3.8	9.6	9.3	15.1
	CV	.22	.18	.24	.18	.20
FCFS	μ	6.5	8.4	7.0	10.1	7.9
	σ	1.3	1.9	1.3	2.4	1.7
	CV	.19	.23	.19	.23	.22
Scen. avg.	μ	10.7	16.4	16.3	33.5	68.6
	σ	2.3	4.8	7.5	16.1	33.5
	CV	.21	.29	.46	.48	.49
Scen. std.	μ	2	3.76	1.87	2.5	1.29
Main effect	F1	<.001	<.001	<.001	<.001	<.001
(p-value)	F2	<.001	<.001	<.001	<.001	<.001
	F3	<.001	<.001	<.001	<.001	<.001

extent to which patients are late once they are over their DT. This information is provided by the weighted DT cost which implicitly considers the tail of the distribution.

The weighted DT cost is based on the idea that patients that passed their DT should be served quickly. The more urgent the patients initial DT category, the fewer days they should be allowed to be served after their DT. Therefore, the cost function is proportional to the number of days a patient is served after his/her DT, but inversely proportional to the patient's initial DT in days. It is defined as:

$$V = \frac{\sum_{i \in I} v_i}{|I|} \tag{4.2}$$

$$v_i = \begin{cases} \frac{7}{dt_i}(s_i - (a_i + dt_i)) & s_i - a_i > dt_i \\ 0 & otherwise \end{cases} \tag{4.3}$$

This cost is zero for patients that are served within their DT. This reflects the idea that from a patient outcome perspective the time when a patient is served does not matter as long as it is before the DT. Moreover, patients that are served after the DT will, at different points in time, eventually be exposed to a similar health risk. For example, a patient with a DT of 1 week who is served 1 week late is associated with the same cost/risk as a patient with a DT of 4 weeks who is served 4 weeks late. The penalties for each day served late for DT categories 4 to 7 are 1, 1/2, 1/4 and 1/8 respectively. No penalty is linked to DT category 8 as they are not given a time limit and their health conditions should generally not worsen over time. A similar idea is used in Riise and Burke [232], who describe a Norwegian setting where violations of due dates are regarded to be one of the measures of a hospital's efficiency.

In reality, surgeries can be performed late for other than scheduling-related reasons. We assume that this is the case for patients who wait for a longer time than 5 times their DT. Those patients are consequently excluded from the cost formula in the performance analysis of the real data.

From Figure 4.2, we conclude that FCFS also performs well with regards to the weighted DT cost. As more DT categories are scheduled FCFS, the average weighted DT cost decreases. This suggests that the benefit of scheduling patients of less urgent DT categories FCFS compensates for the resulting possible delays of more urgent DT category patients. For example, the benefit of providing DT category 5 patients quick access to the OR compensates for the occasionally caused delays of DT category 4 patients.

The benefit of FCFS is the largest if replanning of patients (i.e., bringing them forward in the schedule) is not allowed. Similarly, replanning is able to partly compensate for the benefits of FCFS in the case when applying FCFS is not entirely possible. Consequently, if FCFS is not applicable in reality, it is important to allow for replanning. Replanning should include patients from DT category 8 (Fig. 4.2). This is interesting as DT category 8 does not contribute directly to the DT measure (surgeries of the category have a weight of 0). However, when their surgeries are replanned, they free up future capacity that can be used by surgeries from DT categories that do contribute to the DT cost measure.

One might wonder why FCFS outperforms replanning. FCFS is a regular planning procedure and has to obey the MSS only. Contrary, replanning is more restricted than a regular planning procedure as surgeries cannot be replanned to

Table 4.5 The statistical comparison of the APQ and best-fit strategies using a one-way ANOVA shows that these strategies do not lead to significantly different results in most scenarios.

F1	None			DT4			≤DT5			≤DT6			≤DT7			≤DT8		
F2	E	C	L	E	C	L	E	C	L	E	C	L	E	C	L	E	C	L
Within DT	.89	.33	**.02**	.89	.31	**.04**	.48	.18	.08	.61	.55	.26	.60	.54	.32	1	1	1
Waiting time	.11	.54	**.04**	.11	.26	.17	.32	.28	.25	.65	.73	.31	.58	.85	.66	1	1	1
Weighted DT cost	.16	.23	**.01**	.16	.15	**.03**	.35	.18	.10	.67	.56	.20	.44	.46	.25	1	1	1

Only some p-values are smaller than 0.05 (in bold) for a combination of factors 1 and 2. These are the scenarios where the best-fit strategy performs better than the APQ and where surgeries are scheduled to further dates and thus replanning actions are important. In those setting, the advantage of best-fit compared to the APQ is more apparent.

empty ORs, i.e., a surgery can only be replanned into an OR that has already been assigned to the corresponding surgeon. Replanning surgeries into empty ORs would require the hospital to provide full staffing for entire ORs from one day to another. This is something that we generally would like to avoid.

In Figure 4.2, we see that the APQ does not outperform the best-fit strategy (the triangle typically lays higher than the large gray dot). This is surprising as the APQ ensures that urgent patients in danger of exceeding their DT are replanned first. It is thus tailored to perform well with regards to the weighted DT cost. Further analysis shows that the APQ and best-fit strategies are in most scenarios not performing statistically different with regards to any of the three tested performance measures (Table 4.5) and in some scenarios, the best-fit strategy even outperforms the APQ. The fact that the APQ does not perform better than the best-fit strategy implies that the benefits of replanning are not a result of cost reductions associated with individual patients saved from running late. Instead, it performs well because it saves capacity that otherwise would be wasted. Consequently, the replanning procedure does not need to consider the DT.

The decomposition of the DT cost by DT category (Table 4.6) shows that the highest cost in the model is associated to DT 4, as expected. Interestingly, however, in reality the highest cost is associated to DT category 5. This is unexpected and shows that in reality DT category 4 is efficiently handled. It may also mean that DT category 4 patients might be in certain situations overly prioritized, resulting in exaggerated delays of patients of DT categories 5 and 6. Improving

Table 4.6 Weighted DT cost decomposed by DT category.

		DT4	DT5	DT6	DT7
Real	μ	1.4	3.2	3.1	1.8
	σ	.81	.77	.69	.83
	CV	.60	.24	.22	.48
FCFS	μ	1.6	.60	.02	.00
	σ	.81	.61	.05	.01
	CV	.51	1.0	2.0	6.6
Scen. avg.	μ	3.0	2.7	.45	.09
	σ	.89	1.3	.25	.07
	CV	.29	.47	.56	.81
Scen. std.	μ	1.23	1.1	0.25	0.1
Main effect	F1	<.001	<.001	<.001	<.001
(p-value)	F2	<.001	<.001	<.001	<.001
	F3	<.001	<.001	<.001	<.001

the scheduling of those two categories might therefore lead to the largest benefits for the hospital.

Table 4.6 also shows that DT category 4, in absolute terms, is better handled in reality than in any simulated scenario including FCFS. It seems that some surgeons may always keep some slack capacity reserved for DT 4. The occasional capacity loss might then translate into decreased service levels for DT 5 and 6 patients.

4.2.5 Discipline-specific insights

In Table 4.7 we highlight some of the discipline-specific aspects of the results. As the table shows, the FCFS strategy always performs better than the average scenario (FCFS μ is always better than scen. μ). Generally, FCFS will also give better results than what is currently measured at the hospital. This does not necessarily mean that the FCFS strategy, if implemented, would perform necessarily best in reality as there could be important discipline-specific constraints. Nevertheless, it is an indication that it could generally be beneficial for disciplines to schedule patients to closer dates and not to leave any capacity unused.

Table 4.7 Results for each discipline.

		GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	∪
DT score		.23	.02	.36	.46	.41	.34	.59	.44	.79	.24	.35	.31	.16	.41
Duration est.		2.8	2.1	2.1	5.2	3.8	1.9	2.6	3.6	2.1	1.7	2.6	3.2	2.9	2.7
CV arrival caseload		.30	.70	.28	.36	.38	.40	.50	.37	.23	.27	.44	.60	.29	.17
Slack %		8%	87%	21%	19%	8%	22%	31%	7%	6%	9%	47%	14%	3%	20%
DT offset (%)	real	52.6	-	60.8	69.3	55.3	71.7	93.8	51.0	86.3	47.1	87.4	29.9	51.3	65.6
	FCFS	55.0	-	99.2	72.9	86.6	84.5	89.0	91.0	86.6	95.7	92.1	65.4	95.2	86.3
	Scen. avg.	37.7	-	93.5	63.9	76.3	60.5	76.3	76.9	69.8	57.5	83.1	54.1	74.0	71.5
Waiting time	real	83.9	19.5	40.1	31.9	37.2	16.0	44.4	31.1	12.6	35.5	27.6	74.2	57.4	38.2
	FCFS	15.1	3.4	3.9	13.0	8.2	6.9	4.9	6.9	4.9	5.0	7.3	20.3	9.5	7.8
	Scen. avg.	50.9	62.5	17.8	25.4	20.4	39.6	24.6	21.9	17.4	50.5	23.6	42.4	44.1	30.8
Weighted DT	real	2.6	-	2.8	2.1	3.7	1.9	.20	3.3	.86	4.3	.80	5.4	3.2	2.4
	FCFS	2.1	-	.02	2.4	.57	.49	.46	.29	.35	.09	.37	2.9	.15	.71
	Scen. avg.	5.6	-	.26	3.2	1.5	3.2	1.4	1.0	1.3	3.6	1.1	3.7	1.8	1.9

The performance of individual disciplines can be explained by many factors, such as the arrival caseload variability and the amount of slack capacity. DT-related performance measures for Tx cannot be interpreted as most of their patients are from DT category 8, which is not considered to have a deadline.

Comparing results from the model with reality, we see that a discipline where the performance difference is large is ABD. As the results in Table 4.7 suggest, the discipline performs well in the model but less good in reality. ABD should theoretically be able to handle its patient load very well. This is the case as its arrival caseload is stable (low CV value), the estimated durations are short and little variable, its surgery urgency mix is low (low DT score) and it seems to have enough slack capacity.

One of the reasons why ABD might perform worse than expected could relate to the fact that they accommodate a large amount of non-elective patients. Some surgeons might therefore be more wary of fully utilizing their available capacity and instead leave more slack.

A discipline that seems to have difficulties accommodating its surgeries (in reality and in the model) is MKA. The discipline's major problem factor seems to be the highly unstable arrival caseload (CV is 0.6). This explains the low amount of patients that are served within their DT both in reality and in the model. They could perform better if strategies would be in place that allow them to flexibly control their weekly number of ORs. This might help them to be better equipped for weeks with high loads.

4.3 Discussion

One of the general trends that are observable in the results is that the effectiveness of scheduling factors in utilizing OR capacity will determine how good it performs. This is shown as, firstly, FCFS which is a strategy that disregards the DT but ensures good use of OR capacity, performs very well. Secondly, replanning using the best-fit method performs not worse than the APQ method. This might indicate that avoiding the waste of OR capacity (goal of the best-fit method) is more important than saving individual patients from going over their DT (goal of the APQ). Thirdly, as shown by Figure 4.4, scheduling methods that perform well also yield the highest amount of average next day free capacity. A high amount of average next day free capacity means that there is little blockage and thus OR capacity is used efficiently.

Identifying the efficient use of OR capacity as the dominant driver of patient-related performance also explains some of the seemingly counterintuitive results we got. We noticed that including higher DT categories to be served FCFS does not hinder fast service of lower DT categories, e.g., scheduling DT category 5 FCFS does generally not result in fewer patients from DT category 4 served next day. In reality, there may be cases when DT category 5 patients do hinder quick service of DT category 4 patients. However, this is counterbalanced by the fact that serving DT category 5 FCFS results in an improved use of OR capacity that in return also benefits most DT category 4 patients in the long term.

4.4 Conclusion

We tested the one-step strategy, a method where patients are scheduled to a final surgery date during their consultation session, with the following three factors: FCFS, surgery postponement and replanning to next day. We have shown that all these three tested factors do not influence OR-related performed measures from a practical perspective, but have an effect on patient-related performance measures.

We have shown that in order to serve a large number of patients within their

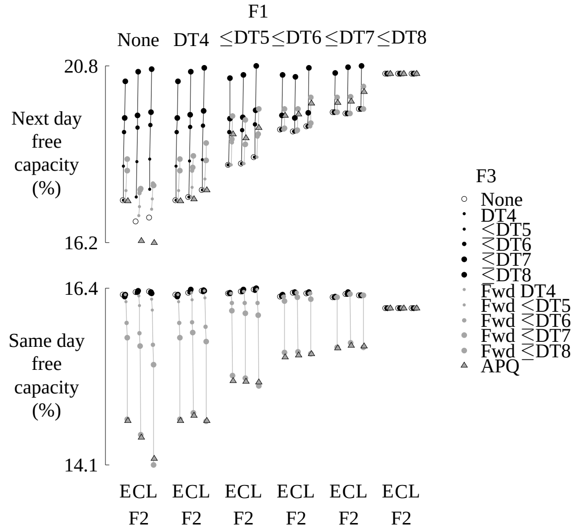


Fig. 4.4 The free capacity on the next day and the same day. The percentage of next day free capacity is observed each day in the morning and relates to the amount of unplanned capacity of the following day before any replanning was done. Same day free capacity relates to the amount of unplanned capacity of the same day in planned open ORs. Same day free capacity, as patients cannot be planned for the same day, corresponds to the day's final amount of unplanned free capacity. However, this capacity might still be used by non-electives or OR-reassigned patients.

DT it is more important to focus on the efficient use of OR capacities than on patient priorities. A strategy that makes good use of OR capacities and therefore performs well is FCFS. If it is not possible to serve patients FCFS, it is important to allow for patient replanning. It is best to replan patients that best fill out free next day OR capacity (best-fit method) instead of focusing on replanning high priority patients. Postponing less urgent surgeries does not benefit DT-related performance measures and only increases patient waiting times.

Chapter 5

Two-step strategy

At the University Hospital Leuven patients who need surgery are assigned directly to an OR and a date. At the hospital, they consider to switch from this one-step scheduling procedure to a two-step procedure. Instead of assigning patients directly to a surgery date and OR, they will be assigned to a week first. This means that a second step is required, where for all the patients that are assigned to a given week a suitable OR and weekday must be determined. The advantage of the two-step procedure is that the second part of the procedure, the within-week scheduling, can be done just before the start of each week (e.g., Thursday or Friday of the preceding week). In other words, a large part of the scheduling decisions can be postponed to a time point close to the surgery date.

One of the main advantages of using the two-step scheduling strategy over the one-step scheduling strategy is that a static component (the within-week scheduling step) is introduced into the scheduling framework. The method also has some drawbacks. For example, the staff and the patient only gets to know next week's exact surgery schedule on the Friday of the current week. In this chapter we investigate the implications of switching to such a two-step scheduling procedure for the hospital.

Unlike in the one-step procedure, the assignment of surgeons to ORs of their discipline is not trivial in the two-step procedure. One of the questions that arise is whether to assign surgeons to ORs (1) in advance (i.e., when the MSS is created), (2) once their patient load for a given week becomes (partly) known or (3) entirely during the second step (i.e., make it part of the within-week scheduling step). Additionally, one also has to determine the circumstances under which surgeons can share certain ORs (e.g., surgeon A occupies a morning block and surgeon B occupies an afternoon block).

In the model, we chose option 3 and therefore assign surgeons to ORs during the within-week scheduling step, that is, once the surgeries to be scheduled for each surgeon become known. Therefore, the available OR days from the MSS are assigned to surgeons based on their exact known surgery load for that given week.

The within-week scheduling step consists of two-steps (not to be confused with the two-step scheduling strategy). In the first step, ORs are assigned to surgeons (surgeon-assignment), and in the second step, individual surgeries are assigned to these ORs (surgery-assignment). Surgeons can only be assigned to ORs that are reserved for their discipline (MSS), while surgeries can only be assigned to ORs that are assigned to their surgeons.

During the surgeon-assignment step, surgeons are iteratively assigned an OR based on their requested capacity (sum of their surgery durations booked for that particular week). First, the surgeon who requests the most capacity is selected and assigned an OR from the weekday (Mon-Fri) with the highest number of free ORs. This decreases the capacity requested by the surgeon by 9 hours (this is the capacity of the OR). The surgeon with the remaining largest capacity request is selected in the next iteration and assigned one of the remaining ORs. Surgeons can only be assigned one OR for a given weekday, therefore they can be assigned a maximum of five ORs a week. The iterative procedure continues until the capacity requests of all surgeons are fulfilled or until there are no ORs left to assign. If a surgeon is not assigned a single OR, then his/her surgeries will be assigned to an OR of another surgeon from the discipline.

Table 5.1 Tested two-step scheduling factors.

	Factor	Values
F1	Step 1: Protection Levels	Off, baseline, break-point
F2	Step 2: Within-week scheduling	WFit, WFit DT
F3	Push	None, DT 4 5, all (DT 4 to 8)

5.1 Factors

We tested three factors: protection levels, DT driven within-week scheduling and a patient push mechanism.

Learning from the results of the one-step strategy, we decided to avoid postponing surgeries and therefore follow FCFS during the to-week scheduling step. Therefore, surgeries will always be assigned to the earliest week possible given the used protection levels.

5.1.1 Factor 1: Protection levels

We are faced with a dynamic scheduling problem and therefore the exact time when a future urgent patient arrives is unknown in advance. In our setting, this can be a problem in situations where a large number of DT 4 or 5 patients arrive while there is no capacity left in the next one or two weeks to accommodate them. Protection levels ensure that some OR capacity remains most of the time reserved (i.e., protected) for such urgent patients.

As capacity is reserved for each DT, patients are postponed in case the capacity reservation requirements for any more urgent DT category makes that necessary. Protection levels are nested and therefore capacity reserved for a given DT category can always be used by a more urgent DT category (e.g., capacity reserved for DT 6 can always be used by a DT 4 or 5 patient. Protection levels are not set for DT category 8 as it contains the least urgent patients and therefore there is no DT category that it could need to have capacity protected from.

A patient can only be scheduled for a certain week if the difference between the free OR capacity and the sum of the protected capacities for more urgent DT categories is larger than the surgery's estimated duration. We set the amount of weekly capacity protected for each DT category equal to the average amount of weekly capacity needed for that DT category. This is the multiplication of the expected number of weekly patients and their estimated average surgery duration.

We use protection levels in two ways. First, with a baseline strategy and, second, with a break-point strategy. The two strategies differ in the way capacities are logged. Using the baseline strategy, surgeries will always consume capacity that has been protected for their DT category. Using the break-point strategy, this will only be the case for surgeries that were scheduled to a date before the break-point. If scheduled after the break-point, unprotected capacity is used whenever possible. The break-point is set equal to the surgery's DT (e.g., for DT 5 it is 2 weeks). Note, that the break-point strategy will always keep the same or more capacity blocked in comparison with the baseline strategy.

5.1.2 Factor 2: Within-week scheduling step

The second factor is used to test whether it is beneficial to consider the DT during the within-week scheduling step. We tested two heuristics: WFit and WFit DT. The former strategy levels OR occupancy, whereas the latter additionally considers the DT of patients. In the description of both strategies, we will assume that the ORs have already been assigned to surgeons. Moreover, surgeries of surgeons without an OR are scheduled last and can be assigned to any OR from the surgeon's discipline.

The name WFit is an abbreviation of the term worst-fit from the memory management literature and is conceptually a strategy that works the reversed way of best-fit. The aim of WFit is to create a schedule where ORs have a similar occupancy. Practically speaking, this translates into a strategy where a surgery is always assigned to the OR with the most leftover capacity. Therefore, patients are first always assigned to empty ORs. After each OR is occupied by at least one patient, the next patients are iteratively assigned to the OR that has the most remaining capacity left. The algorithm starts with the longest surgery and ends with the shortest surgery, therefore the schedule in each OR will also start with longer and end with shorter surgeries. As this is not desired, surgeries in each OR are shuffled.

An extension of WFit is WFit DT where also the DT of patients is considered. This algorithm starts by assigning patients to one of six groups. The first group contains those patients that are already late (i.e., even if scheduled for Monday, they are served after their DT). The second, third, fourth and fifth group contain patients that have to be served the latest by Monday, Tuesday, Wednesday and Thursday respectively. The sixth group contains those patients who will remain scheduled within their DT even if scheduled for Friday. Within each of the six groups, surgeries are processed from long to short. The algorithm first schedules group one. The patients in the first group are scheduled for Monday first. A patient that is assigned to a date is removed from its group. At this stage, ORs are not allowed to be overbooked, that is, the sum of the estimated surgery durations assigned to an OR has to be smaller than or equal to nine hours. If all patients in the first group are scheduled or, alternatively, there is not enough capacity on Monday left, then the algorithm enters the second stage. In the second stage the remaining patients from the first group are merged with the patients from the second group. The patients in the newly formed group are scheduled for both Monday and Tuesday ORs. As before, ORs are not allowed to be overbooked. The same procedure continues until patients from all six groups are included. At the last stage patients can be scheduled to all weekdays. Surgeries left unassigned are overbooked and are therefore divided among all ORs using the basic WFit algorithm.

5.1.3 Factor 3: Push

The third factor is used to test a push mechanism that allows patients to be assigned to a date that is in the same week as they arrived. If patients are pushed into the current week, then they skip both steps of the scheduling procedure.

Using the push factor, a patient that arrives Wednesday can be served on Thursday or Friday of the same week. This would normally not be possible as the first step of the two-step procedure requires patients to be assigned to a complete week (i.e., the succeeding Monday would be the closest possible surgery date).

In reality, it is better to give the patient some time to prepare for their surgery if the health condition of the patient allows it. We therefore also tested the push 4 5 factor, which only allows DT 4 and 5 patients to be pushed into the current week and therefore DT 6, 7 and 8 patients can only be scheduled with the two-step procedure.

5.2 Results

The factors of the two-step strategy have fewer realizations than the factors of the one-step strategy. As a result, the constructed full factorial design consists of 18 scenarios only. Two-step scenarios are evaluated on both OR and patient-related performance measures.

5.2.1 OR-related performance measures

Most OR-related performance measures are only to a very limited extent affected by the two-step factors. This is shown by the fact that the standard deviation between different scenarios is very small (Table 5.2). For example, overtime is only to a very limited degree affected by the chosen value for each factor (Fig. 5.2). In Section 4.2.1, we made similar conclusions for the one-step factors.

An exception to this rule is the cancellation rate (Fig. 5.1). The two factors that exert the greatest influence on the cancellation rate are the within-week scheduling factor and the push factor. The fact that the cancellation rate increases if the within-week scheduling step considers the DT is understandable as this can cause an unbalanced weekly schedule (e.g., if many patients have to be served by Monday, then the ORs on Monday will be fully utilized, while the ORs on the other days of the week might show a lower utilization).

Table 5.2 Two-step strategy: OR-related performance measures.

		Utiliz-ation %	Over-time %	Under-time %	Open OR Daily mean	Throughput Daily mean	Reassign-ed %	Cancel-led %
Real	μ	83.4	13.6	6.0	18.0	47.9	13.1	3.4
	σ	1.3	1.6	1.2	1.5	4.2	2.8	1.1
	CV	.02	.11	.19	.08	.09	.21	.33
Full	μ	88.2	15.8	4.7	16.7	48.1	13.9	4.7
	σ	1.6	2.0	1.2	.26	2.1	1.9	1.7
	CV	.02	.13	.26	.02	.04	.13	.37
Scen. avg.	μ	88.0	15.8	5.0	16.7	48.1	13.6	5.0
	σ	.19	.21	.34	.03	.01	.41	.52
	CV	.00	.01	.07	.00	.00	.03	.10
Scen. std.	μ	1.65	2.07	1.39	0.25	2.07	1.89	2.01
Main effect (p-value)	F1	<.001	.27	<.001	<.001	.99	<.001	<.001
	F2	.05	<.001	.06	<.001	100	<.001	<.001
	F3	.95	<.001	.13	<.001	.97	<.001	<.001

The full strategy represents the scenario where all factors are used to the fullest, that is, protection levels are used with the breakpoint strategy, the within-week scheduling step considers the DT and push is applied to all DT categories.

It is more difficult to explain why pushing patients into the fixed weekly schedule results in a lower cancellation rate. Pushing patients into free capacity results in a more dense schedule, increasing the probability of cancellations. By now it should be clear that the reason for a reduced cancellation rate stems from the fact that applying the push factor allows to use capacity more efficiently resulting in less cancellations in the long run.

Since the majority of OR-related performance measures, from a practical perspective, do not affect scheduling factors, we concentrate on patient-related performance measures (Table 5.3). We investigate in particular: the percentage of patients served within their DT, the average patient waiting time and the weighted DT cost (Fig. 5.3).

5.2.2 Percentage of patients served within DT

Large improvements can be achieved using DT driven within-week scheduling as WFit DT scenarios consistently outperformed WFit scenarios (Fig. 5.3). Be-

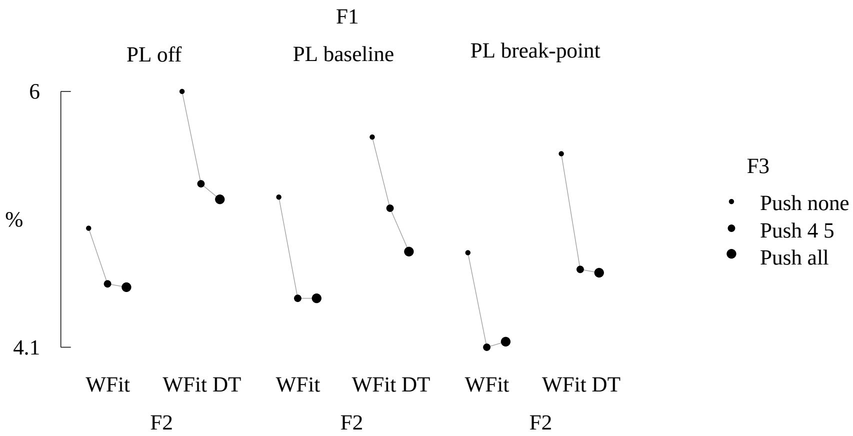


Fig. 5.1 Two-step strategy: the cancellation rate depends on the chosen scheduling strategy.

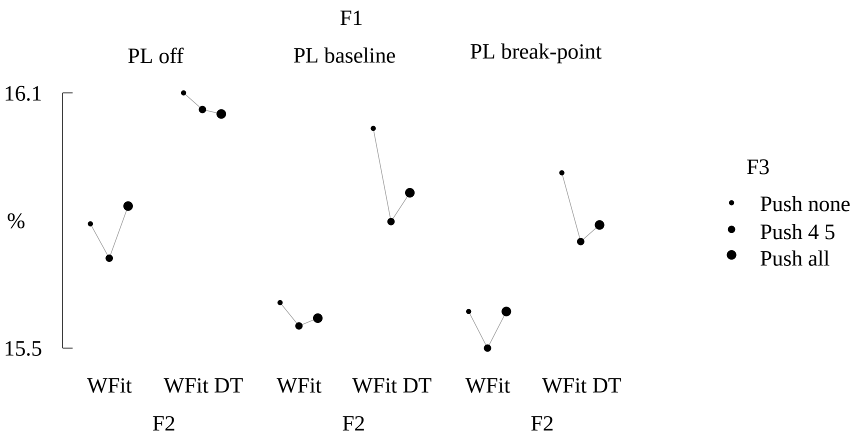


Fig. 5.2 Two-step strategy: the amount of overtime, from a practical perspective, does not depend on the considered two-step factors (the minimum and maximum values are very close).

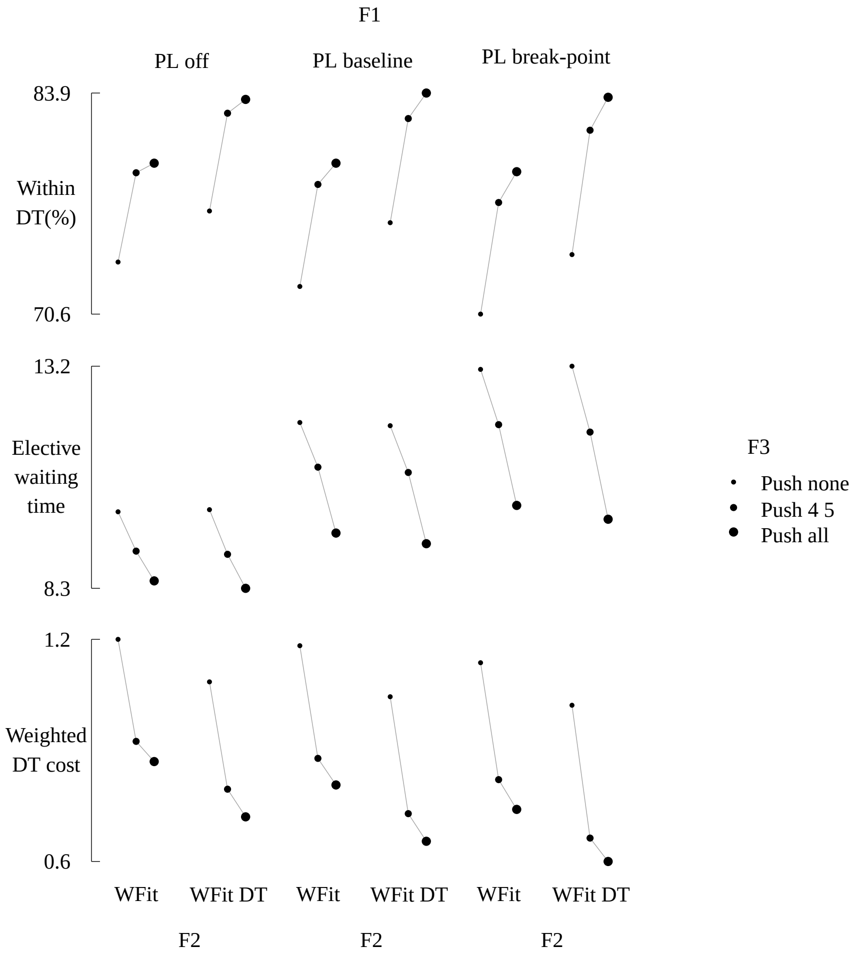


Fig. 5.3 Two-step strategy: patient-related performance measures for the two-step procedure.

Table 5.3 Two-step strategy: patient-related performance measures.

		DT offset (%)	Elective waiting time	Weighted DT cost
Real	μ	65.6	38.2	2.4
	σ	4.4	4.0	.43
	CV	.07	.11	.18
Full	μ	83.6	9.8	.56
	σ	5.5	1.0	.32
	CV	.07	.11	.57
Scen. avg.	μ	78.6	10.5	.86
	σ	4.1	1.5	.21
	CV	.05	.15	.24
Scen. std.	μ	5.34	1.09	0.41
Main effect (p-value)	F1	<.001	<.001	<.001
	F2	<.001	<.001	<.001
	F3	<.001	<.001	<.001

All factors have a significant main effect on all three performance measures.

sides applying the WFit DT factor, it is also important to apply the push factor. Interestingly, the performance gains can almost entirely be realized by applying the push factor for DT 4 and 5 only. This keeps the total number of patients who are pushed into the schedule much lower. Contrary to the WFit DT and the push factor, it is not beneficial to use protection levels.

5.2.3 Patient waiting time

The only factor that has a positive thus decreasing effect on waiting time is the push factor (Fig. 5.3). Factor WFit DT has no real impact on waiting time, whereas protection levels have a negative impact and thus increase the average patient waiting time. This is to be expected as the latter two factors were specifically developed to improve DT-related performance measures and were not designed to decrease patient waiting time.

Investigating the waiting time of each DT category individually gives a more accurate picture of the effects of the three factors (Fig. 5.4 and Table 5.4). We see that using protection levels results in a small benefit for DT 4 patients. However,

Table 5.4 Two-step strategy: waiting time by DT category.

		DT4	DT5	DT6	DT7	DT8
Real	μ	8.1	21.6	40.1	51.7	75.1
	σ	1.8	3.8	9.6	9.3	15.1
	CV	.22	.18	.24	.18	.20
Full	μ	7.3	8.6	10.0	15.9	10.8
	σ	1.0	1.1	1.5	3.7	1.8
	CV	.14	.12	.15	.23	.17
Scen. avg.	μ	8.3	9.2	10.5	16.0	11.6
	σ	.79	.88	1.5	4.2	2.6
	CV	.10	.10	.14	.27	.22
Scen. std.	μ	1.25	1.18	1.38	3.36	1.8
Main effect	F1	<.001	<.001	<.001	<.001	<.001
(p-value)	F2	<.001	.04	.78	.26	.61
	F3	<.001	<.001	<.001	<.001	<.001

Interestingly, larger scenario averages are observed for DT category 7 than for DT category 8. This can be explained by the fact that the distribution of DT categories is heterogeneous across disciplines. The fact that DT category 7 patients experience longer waiting times means that the disciplines that were generally observing longer waiting times also had more DT category 7 patients.

for the remaining DT categories we observe increasing waiting times. Applying the WFit DT factor seriously benefits DT 4 patients, to a lesser extent DT 5 patients and has no significant effect on the remaining categories (Table 5.4). The only factor that reduces the waiting times for all DT categories is the push factor. As Figure 5.4 shows, restricting the push factor to DT 4 and 5 only will still benefit higher DT category patients.

5.2.4 Weighted DT cost

Most of our results are based on interaction plots, which provides the exact value of a performance measure for each combination of factors (these are the scenarios). One of the motivations to use interaction plots is that most factors and performance measures show significant interaction effects and thus investigating main effects only could lead to misleading conclusions. The weighted DT cost is an outlier amongst the tested patient-related performance measures as

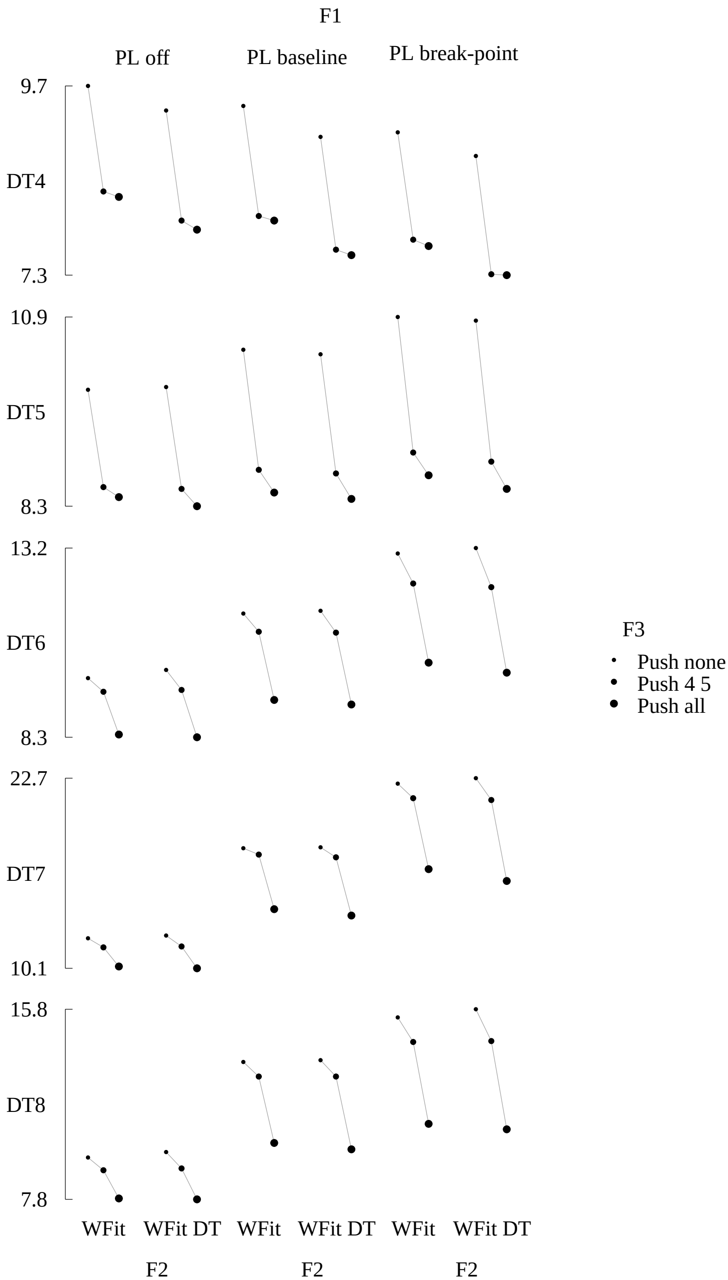


Fig. 5.4 Two-step strategy: the waiting time of each elective DT category.

Table 5.5 Two-step strategy: weighted DT cost decomposed by DT category.

		DT4	DT5	DT6	DT7
Real	μ	1.4	3.2	3.1	1.8
	σ	.81	.77	.69	.83
	CV	.60	.24	.22	.48
Full	μ	1.4	.25	.04	.00
	σ	.86	.22	.06	.00
	CV	.61	.88	1.6	6.8
Scen. avg.	μ	2.2	.35	.06	.01
	σ	.55	.13	.05	.01
	CV	.24	.36	.90	1.4
Scen. std.	μ	1.1	0.28	0.08	0.02
Main effect	F1	<.001	<.001	<.001	<.001
(p-value)	F2	<.001	<.001	.04	.44
	F3	<.001	<.001	<.001	<.001

significant interaction effects are not present. We therefore also look at the main effect of factors on this performance measure.

Figure 5.5 shows that all factors decrease the weighted DT cost. The push factor provides the greatest benefit, while protection levels and the WFit DT factor provide smaller improvements.

Only judging from the main effect, one could conclude that protection levels are always beneficial. Looking at the interaction plot (Fig. 5.6, top right graph), one sees that this is only partly true and that protection levels are best used in combination with the push factor as otherwise the gains are minimal. To a lesser extent, a similar conclusion can be made for factor WFit DT.

Investigating the differences between the cost components of the DT cost (Table 5.5) shows that the largest contributors are DT 4 patients (1.4). A much smaller cost component is due to late DT 5 patients (0.25) and virtually no costs occur from DT category 6 and 7 patients. Consequently, DT 4 patients will determine the weighted DT cost most and therefore one would expect that factors that benefit DT 4 should thus generally also help to decrease the total weighted DT cost. Indeed, WFit DT, a factor that will often benefit DT 4 patients, does result in an overall lower weighted DT cost (Fig. 5.3).

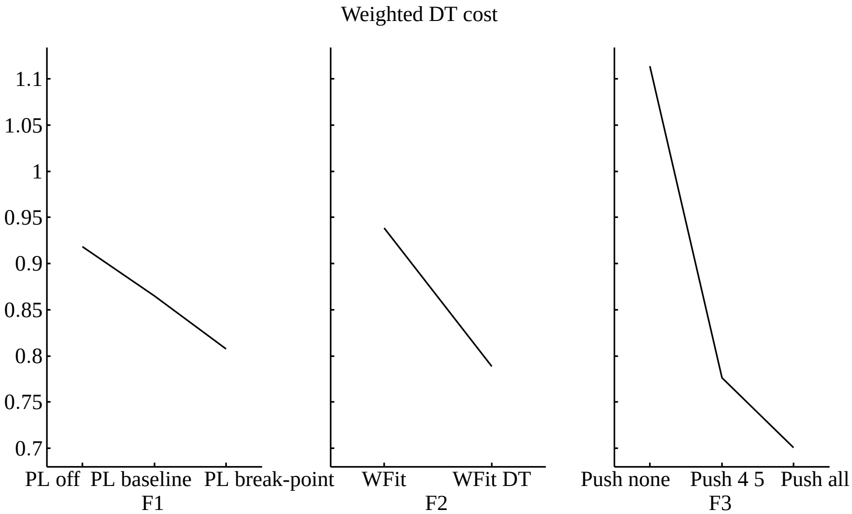


Fig. 5.5 Two-step strategy: main effect of factors on the weighted DT cost. The main effect is the effect of the factor averaging over all other factors.

5.2.5 Discipline-specific insights

For most disciplines, the performance measurements obtained with the full strategy are better than they are in reality (Table 5.6). A discipline for which this is not true is TRH. TRH contains many DT 4 patients and these are difficult to schedule efficiently with the two-step strategy. In the two-step strategy, DT 4 patients are mostly served in the week succeeding the week of their arrival, which often leads to waiting times longer than one week (i.e., they are served after their DT).

Discipline RHK also contains a high ratio of DT 4 patients, but it performs substantially better in the simulation model than TRH. This is because RHK has a large amount of slack capacity and contains, compared to TRH, more patients from less urgent DT categories (Fig. 1.2). Having lower urgency patients can be beneficial during the within-week scheduling step as these can be scheduled towards the end of the week.

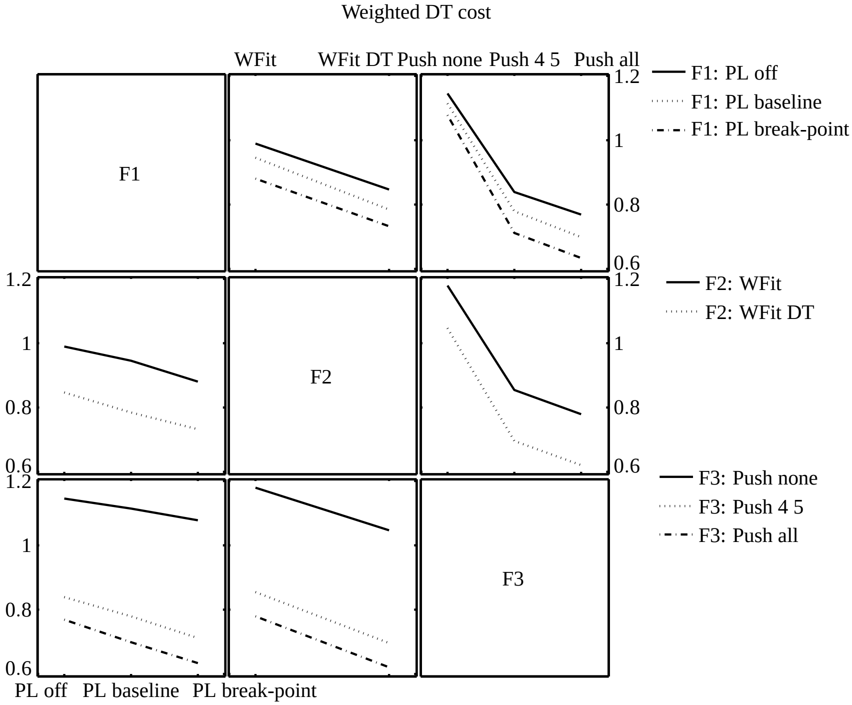


Fig. 5.6 Two-step strategy: interaction effect of factors using the weighted DT cost. The interaction plot shows the effect of one factor given the other two factors.

Table 5.6 Two-step strategy: results for each discipline.

		GYN	Tx	ABD	CAH	NCH	ONC	RHK	THO	TRH	URO	VAT	MKA	NKO	∪
DT score		.23	.02	.36	.46	.41	.34	.59	.44	.79	.24	.35	.31	.16	.41
Duration est.		2.8	2.1	2.1	5.2	3.8	1.9	2.6	3.6	2.1	1.7	2.6	3.2	2.9	2.7
CV arrival caseload		.30	.70	.28	.36	.38	.40	.50	.37	.23	.27	.44	.60	.29	.17
Slack %		8%	87%	21%	19%	8%	22%	31%	7%	6%	9%	47%	14%	3%	20%
DT offset (%)	Real	52.6	-	60.8	69.3	55.3	71.7	93.8	51.0	86.3	47.1	87.4	29.9	51.3	65.6
	Full	94.7	-	97.3	74.2	81.2	89.3	83.6	83.8	61.2	94.7	95.6	78.8	95.3	83.6
	Scen. avg.	88.9	-	95.8	72.4	77.3	85.2	78.5	79.9	49.6	89.5	92.8	75.4	85.7	78.6
Waiting time	Real	83.9	19.5	40.1	31.9	37.2	16.0	44.4	31.1	12.6	35.5	27.6	74.2	57.4	38.2
	Full	9.4	5.6	6.5	16.3	12.6	7.1	6.2	11.5	8.8	6.9	7.2	16.8	16.7	9.8
	Scen. avg.	10.3	6.5	7.0	14.9	13.0	7.7	6.7	11.3	9.1	8.2	8.1	17.4	21.4	10.5
Weighted DT	Real	2.6	-	2.8	2.1	3.7	1.9	.20	3.3	.86	4.3	.80	5.4	3.2	2.4
	Full	.10	-	.05	1.2	.64	.26	.44	.39	1.3	.11	.12	1.2	.11	.56
	Scen. avg.	.24	-	.09	1.7	.97	.41	.65	.60	2.0	.28	.23	1.5	.45	.86

5.3 Discussion

5.3.1 Performing the second stage on Thursday instead of Friday

While the one-step strategy is a purely dynamic scheduling method, the two-step strategy also contains a static scheduling component (the within-week scheduling step). This static component benefits the two-step strategy. Most articles in the inpatient scheduling literature deal with a static setting (Sec. 2.3.3).

Unfortunately, the two-step scheduling strategy also has some disadvantages. Most notably, it leaves the personnel of the hospital with only little time to prepare for surgeries that are scheduled to a day in the beginning of the week. To compensate for this drawback, we also investigate a method where the within-week scheduling step is performed on Thursday afternoon (TH scenario) instead of on Friday afternoon (FR scenario).

In the TH scenario, Friday arrivals are handled differently. In the TH scenario, on Fridays, the next-week schedule is already fixed, while in the FR scenario it is still undetermined. This is a disadvantage of the TH scenario if Friday arrivals need to be served the next week as it then becomes necessary to apply the push factor (Fig. 5.7)).

Because of this disadvantage, it is not surprising that the FR scenario outperforms the TH scenario (Fig. 5.8). The FR scenario has a higher number of patients served within their DT, generally shorter average waiting times, a lower weighted DT cost and less patients pushed into the schedule (Fig. 5.9).

The degree to which the FR scenario performs better depends on the various factors. For example, around 6.5% more patients are served within their DT assuming push is not allowed and around 3.5% if push is allowed.

Restricting the push factor to DT 4 and 5 results in the FR scenario in around 5.5% patients that are pushed in, while in the TH scenario this is around 7.5%. The additional 2% of surgeries pushed into the schedule corresponds to around one extra patient per day over all disciplines and thus is acceptable.

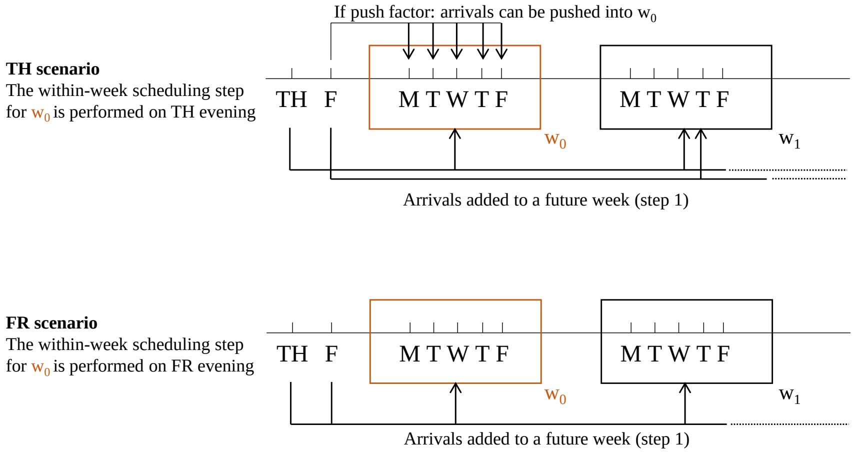


Fig. 5.7 Difference between the TH scenario and the FR scenario. In the TH scenario, the only way how to serve Friday arrivals the next week is using the push factor. In the FR scenario, arrivals don't have to be pushed into the next week as they can be assigned to the next week during the to-week scheduling step.

As we have shown, the TH scenario results in worsened DT-related measures and an increased number of surgeries pushed into the schedule. We think that while the increased number of pushes are acceptable, the worsened DT-related performance measures pose a problem.

Nevertheless, applying the TH scenario can benefit the hospital and the patients as they gain more time to prepare for surgeries. The hospital could also consider a partial solution where the within-week scheduling step for disciplines with few DT 4 patients is executed on Thursday, while for disciplines with a larger DT 4 patient population it is executed on Friday (e.g., for CAH, NCH, ONC, RHK, THO, TRH, URO, VAT, MKA).

5.3.2 Comparison with the one-step strategy

The results for the one-step and two-step strategy can directly be compared against each other as they are retrieved from the same simulation environment (i.e., same non-elective allocation model, same simulation seeds, etc.).

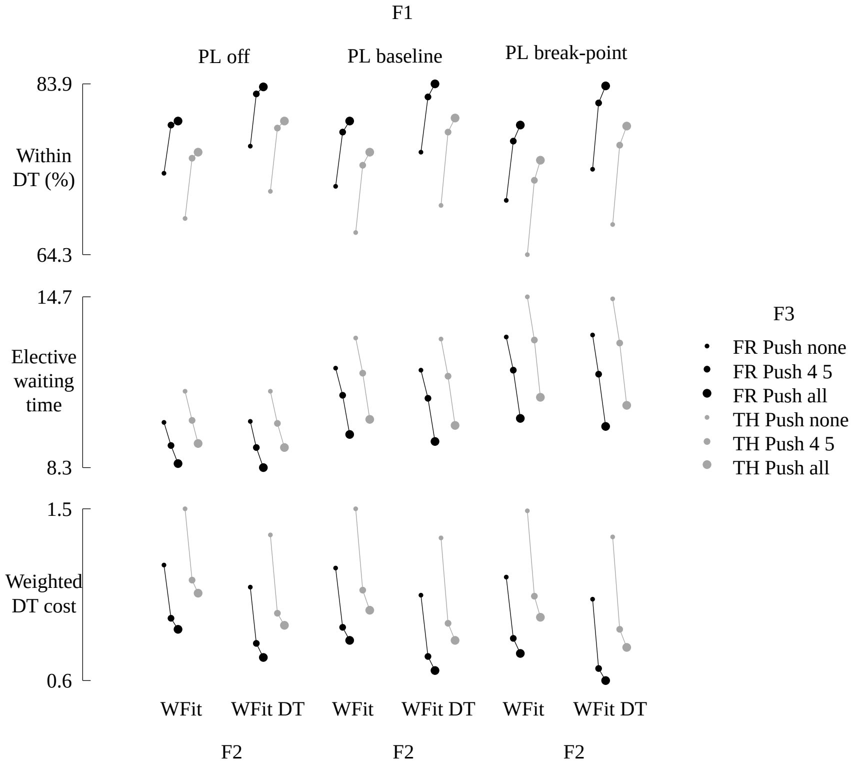


Fig. 5.8 The comparison of patient-related performance measures in the TH and FR scenario.

We compare the strategies based on the FCFS scenario (one-step) and the full scenario (two-step). The former performs better with regards to the percentage of patients served within DT (86.3% versus 83.6%), but worse with regards to the weighted DT cost (.71 versus .56).

Larger differences between the two strategies can also be observed on the discipline level. Based on the weighted DT cost, we see that most disciplines perform better using the two-step strategy (Table 5.6 and 4.7). However, there are exceptions. For example, TRH performs better in the one-step strategy than in the two-step strategy (Table 5.6 and 4.7).

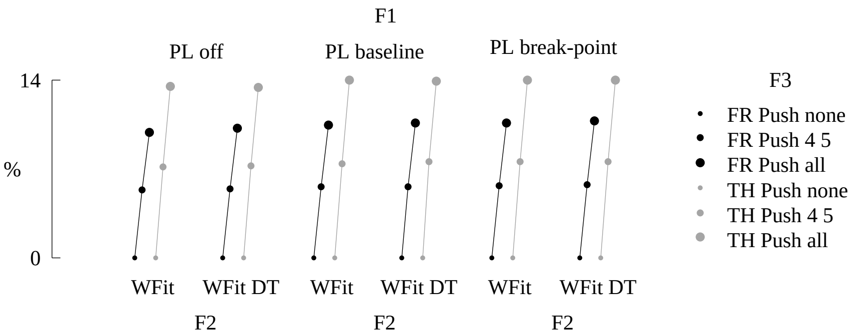


Fig. 5.9 The comparison of the percentage of patients pushed into the schedule in the TH and FR scenario.

There are also disciplines for which the difference between the two strategies is not obvious. For example, ONC and URO are disciplines where the FCFS and full scenarios perform similar, but the scenario averages are better for the two-step strategy.

Disciplines that perform better under the two-step strategy are GYN and MKA. The reason for the good performance of GYN is straightforward: it is a discipline with a very small DT 4 population. Thus the two-step strategy is not burdened with the DT 4 related disadvantage.

The reasons why MKA performs better are less obvious. It is the discipline with the longest surgery durations. Longer surgery durations are generally difficult to efficiently fit into surgery blocks. In cases such as these, using static scheduling, even if only partially as in the case with the two-step strategy, can prove to be beneficial.

5.4 Conclusion

In this chapter we tested the two-step strategy where surgeries are assigned to a week first, and only in a second step to an OR and a weekday. We tested a combination of the following three factors: protection levels, DT driven within-week scheduling and a patient push mechanism.

We have shown that protection levels, a strategy that reserves capacity for high urgency patients, does not perform as expected since it does not decrease the percentage of patients served within the DT, increases patient waiting time and decreases the weighted DT cost only if combined with the push factor.

To consider the DT during the within-week scheduling step was shown to be beneficial. It improves all tested DT-related performance measures, but as it increases the cancellation rate, it comes at a cost. Allowing patients to be pushed into the fixed weekly schedule brings the largest benefits among tested factors as it drastically improves all three tested patient-related performance measures, while it also decreases the cancellation rate.

We have also tested the benefit of performing the within-week scheduling step on Thursday instead of on Friday. We concluded that for disciplines with a large proportion of DT 4 patients this results in a higher weighted DT cost.

Chapter 6

Discussion

6.1 Three points on the literature

In this section we describe some of the observations and conclusions we made with regards to the general OR planning and scheduling literature.

In Section 6.1.1 we discuss how to make a clearer distinction between theory-oriented articles targeting researchers and practice-oriented articles targeting both researchers and practitioners.

In Section 6.1.2 we discuss how some PMs (e.g. overtime) are used universally in articles and why we think more attention has to be paid to selecting specific PMs.

In Section 6.1.3 we discuss points that need to be included in each paper in order to make it easier to situate them in the literature and thus to classify them. Including those points additionally allows readers to determine in an easier way whether the methods or results described in an article are of interest to them.

6.1.1 Clarifying the target group: Researchers or practitioners

In the literature a clearer distinction needs to be made between theory-oriented articles targeting researchers and practice-oriented articles targeting both researchers and practitioners (Table 6.1). Because of publishing reasons articles often address both groups, despite the fact that their actual core contribution is usually only meant for one of those groups. This carries some risks as it overstates those insights that do not result from the main strengths of the paper. This is a problem for both theory- and practice-oriented articles and could be prevented by having a clear distinction between both types of articles concerning their target group and the resulting conclusions. This would also make it easier for readers of both target groups (researchers and practitioners) to confidently identify articles that are relevant for them.

The distinction between theory- and practice-oriented articles starts already in the addressed problem and the research task (Table 6.1). For a theory-oriented paper, the goal is to improve a methodology by solving a specific drawback of it that limits its real-life applicability (e.g., an efficient MP that is able to include various sources of uncertainty), while the goal of a practice-oriented paper is to solve a real-life problem (e.g., an MP that includes all constraints and uses PMs that are relevant for the specific real-life problem). As a consequence, for the former, the collaboration with practitioners is not a prerequisite, while it is essential for the latter one.

It can be a problem if theory-oriented articles target practitioners as they might include managerial conclusions which, without understanding the underlying operations research model, might not be interpreted in the right way by practitioners. As these articles mostly focus on a specific method, they will only include those aspects of the real setting that can be implemented using their method and might also, understandably, overemphasize aspects of the real setting that help them to exemplify a certain advantageous property of their method. This way, those aspects that are important in reality, but cannot be included using the chosen model might be left out. Therefore, it might be beneficial if they direct their research towards other researchers and thus make conclusions mainly on methodological aspects. Naturally, they can still report on preliminary insights from a hypothetical case example as these help to guide other researchers to promising areas where the method's real-life applicability can be put to the test.

Similarly, but perhaps to a lesser extent, it is also a problem if authors of practice-oriented articles overemphasize the role of their model adaptations. Clearly, models are often adapted to the real-life problem setting at hand, but generally those adaptations do not fundamentally improve the methods and therefore will not substantially contribute to the theoretic modeling literature. However, as they can generally choose the best suited method for the problem, they are less restricted by the method's capabilities, which makes it easier for them to focus on including aspects that are important in the real setting.

It is mainly in the conclusion section where the lack of a clear distinction might cause problems (Table 6.1). This is the case for both theory- and practice-oriented articles. For the former this is the case if, for instance, a paper that focuses on a new method mainly concludes on the insights from the testing phase performed on a (perhaps hypothetical) case study instead of drawing conclusions about the modeling adaptations. For the latter this is the case if, for instance, a real problem is solved and the conclusion mainly focuses on detailing the adaptations made to the model instead of on showing the implications for practitioners.

Also the editors of journals can play an important role, since they can ensure that authors are consistent in addressing their target audience (as described in Table 6.1). Additionally, they also ensure that the audience targeted by these authors coincides with the readers of the particular journal. They can ensure this alignment on the one hand by providing adapted publishing incentives for both theory-oriented and practice-oriented articles and on the other hand by clearly positioning the journal.

Table 6.1 Distinction between theory- and practice-oriented articles.

Theory-oriented	Practice-oriented
The **target group** covers	
researchers	researchers and practitioners *E.g., medical staff, hospital managers, policy makers*
The **addressed problem** is	
an operations research method that has drawbacks limiting its real-life applicability. *E.g., a stochastic dynamic program that is not tractable for patient test sets of realistic size*	a real-life OR planning problem that has no efficient solution yet. *E.g., an inefficient surgery rescheduling policy at a case hospital*
The **research task** involves	
identifying important aspects of the method that need to be improved to ensure real-life applicability. *E.g., aspect that can reduce the dimensionality of the state space in the model formulation*	identifying important aspects of the real setting that need to be included into the model to ensure realism. *E.g., factors that trigger rescheduling such as the arrival of an emergency patient*
identifying approaches that can be used to solve the identified drawbacks. *E.g., aggregate the state space*	identifying methods that can be used to solve the problem at hand. *E.g., an advanced MP approach able to include various personnel constraints*
using objectives and assumptions that are relevant in the context, but are possibly motivated by the literature. This does not require collaboration with practitioners. *E.g., a trade-off between overtime and waiting time, Poisson arrival distribution*	using objectives and assumptions that are realistic and importantly, motivated by the setting. This requires collaboration with practitioners. *E.g., a trade-off between cancellations and overtime, only reschedule to ORs with suitable equipment*
The **findings** include	
the method improvement itself (e.g., relaxation of an assumption). *E.g., the model can now solve datasets with up to 10 ORs, where before this was limited to 5 ORs*	confirming the applicability of a method to the problem at hand. *E.g., the algorithm provides an efficient rescheduling mechanism and has a reasonable running time*
results on the testing phase, which only showcases the capabilities of the improved method using a (hypothetical) example, supported by a scenario analysis. *E.g., based on a hospital with 10 ORs, 7 disciplines and 6 surgery types, the method created an optimal schedule*	results on the testing phase needed to make conclusions. The results are supported by an extensive data analysis. *E.g., the developed policy reduces overtime and the number of cancellations by 10% and 3% respectively*
results on the computational performance of the improved method. *E.g., the model solves all tested scenarios to optimality in less than one hour*	results on the application of the proposed solution (if used in practice). *E.g., adopting the derived decision rules, the staff experienced less overtime and fewer equipment conflicts*
The **conclusions**	
discuss the idea behind the model advancement that led to the beneficial properties of the model. *E.g., aggregating the state space drastically reduces its dimensionality*	discuss the implications the tested decision rules or policies have on the case hospital. For algorithms, they discuss the derived rules or policies. *E.g., analyzing the results of the algorithm showed that rescheduling the patient with the longest duration first results in the best trade-off*
identify promising real-life examples to test the method's applicability. *E.g., the method can now not only solve strategic problems, but also problems that need to be solved daily*	identify promising ways to improve the used method's real-life applicability, i.e., show its limitations. *E.g., once the number of ORs increases, the run time of the algorithm increases drastically*
discuss those insights that can be generalized or used for improvements on other methods. *E.g., Erasing from the memory that part of the state space that will not be used anymore by the algorithm allows to solve problems of realistic size. In other methods this logic of explicitly tagging and erasing data that will not be used anymore can also be used*	discuss those insights that can be generalized or used by other hospitals. *E.g., rescheduling decreases overtime only if OR closing times are flexible and not if they are fixed.* if possible, include comments from practitioners. *E.g., the staff suggests including personnel preferences*

6.1.2 Clarifying the objective

We observed that some PMs (e.g., overtime) are used in articles irrespective of the tackled problem setting (i.e., the combination of a decision and an assignment level). In order to better understand how they depend on the problem setting we test their dependency using a Fisher test (Table 6.2). Unlike the Chi-square test, this test can also be used with low sample sizes.

The results show that 5 out of 9 PMs are not used in a setting-specific way. This is surprising, since we would generally expect that for a given problem setting, given PMs apply. Importantly, for this analysis we simplified the problem setting. Although articles often cover more than one decision and assignment level, we will only look at pairs. For example, the assignment of ‘patient’ to ‘day’ and ‘room’ is split into two cases: ‘patient’ to ‘day’ and ‘patient’ to ‘room’.

As many PMs are used independently of the problem setting, one might wonder whether PMs are generally used in an appropriate way. We think that this is not always the case and argue that it is important to choose PMs carefully, keeping in mind the following two steps.

First, choose the PM that is of practical relevance to the real setting. This means that it captures the most important objective(s) according to the stakeholders, not (only) according to the research community. For example, the average waiting time is an important criterion in many settings. However, if a diverse patient population is assumed, it might not suffice to decrease the average waiting time. For instance, a patient that is waiting for a hip replacement and a patient with metastatic cancer clearly do not exhibit the same urgency. A scheduling method that cuts the average patient waiting time might slightly benefit the former patient category, but seriously harm the latter one, which might not be in line with the hospital targets. The chosen PM thus needs to be the result of a thorough discussion with the stakeholders on the desired outcome or improvement.

Second, once appropriate PMs have been defined, check whether for each PM the associated mechanisms are included in the model. This means that the model needs to contain those mechanisms that principally determine the value of the PM. For example, for a scheduling algorithm in an inpatient setting where over-

Table 6.2 Selecting appropriate PMs should not be done based on their popularity in the literature.

<table>
<tr><th>Setting</th><th colspan="9">Performance measures: observed|expected</th><th>Count</th></tr>
<tr><th></th><th>Patient waiting</th><th>Over-util. OR</th><th>Under-util. OR</th><th>Utiliz. OR</th><th>Through-put</th><th>Prefe-rence</th><th>Finan-cial</th><th>Make-span</th><th>Defer-ral</th><th></th></tr>
<tr><td>Disc-Day</td><td>7|6.3</td><td>7|11.5</td><td>4|4.4</td><td>4|4.6</td><td>6|3.7</td><td>3|3.9</td><td>4|2.8</td><td>0|2.4</td><td>3|3.4</td><td>22</td></tr>
<tr><td>Disc-Time</td><td>4|3.4</td><td>6|6.3</td><td>0|2.4</td><td>5|2.5</td><td>6|2</td><td>2|2.1</td><td>2|1.5</td><td>0|1.3</td><td>3|1.9</td><td>12</td></tr>
<tr><td>Disc-Room</td><td>6|5.4</td><td>6|9.9</td><td>4|3.8</td><td>4|3.9</td><td>6|3.2</td><td>4|3.4</td><td>1|2.4</td><td>0|2.1</td><td>3|3</td><td>19</td></tr>
<tr><td>Disc-Cap</td><td>7|6</td><td>8|11</td><td>3|4.2</td><td>10|4.4</td><td>9|3.6</td><td>3|3.7</td><td>4|2.6</td><td>0|2.3</td><td>6|3.3</td><td>21</td></tr>
<tr><td>Surg-Day</td><td>3|4.3</td><td>5|7.8</td><td>2|3</td><td>4|3.1</td><td>2|2.6</td><td>1|2.7</td><td>2|1.9</td><td>1|1.6</td><td>4|2.3</td><td>15</td></tr>
<tr><td>Surg-Time</td><td>2|3.4</td><td>6|6.3</td><td>0|2.4</td><td>2|2.5</td><td>1|2</td><td>1|2.1</td><td>2|1.5</td><td>1|1.3</td><td>3|1.9</td><td>12</td></tr>
<tr><td>Surg-Room</td><td>2|4.3</td><td>7|7.8</td><td>2|3</td><td>3|3.1</td><td>2|2.6</td><td>1|2.7</td><td>4|1.9</td><td>2|1.6</td><td>2|2.3</td><td>15</td></tr>
<tr><td>Surg-Cap</td><td>3|4.3</td><td>6|7.8</td><td>2|3</td><td>3|3.1</td><td>1|2.6</td><td>1|2.7</td><td>8|1.9</td><td>1|1.6</td><td>3|2.3</td><td>15</td></tr>
<tr><td>Pat-Day</td><td>28|25.2</td><td>53|45.9</td><td>27|17.4</td><td>16|18.3</td><td>12|15</td><td>23|15.6</td><td>8|11.1</td><td>6|9.6</td><td>14|13.7</td><td>88</td></tr>
<tr><td>Pat-Time</td><td>21|21.8</td><td>41|39.7</td><td>10|15.1</td><td>12|15.8</td><td>11|12.9</td><td>10|13.5</td><td>4|9.6</td><td>20|8.3</td><td>8|11.9</td><td>76</td></tr>
<tr><td>Pat-Room</td><td>25|27</td><td>56|49.1</td><td>24|18.6</td><td>18|19.5</td><td>12|16</td><td>19|16.7</td><td>4|11.8</td><td>15|10.3</td><td>6|14.7</td><td>94</td></tr>
<tr><td>Pat-Cap</td><td>15|11.5</td><td>23|20.9</td><td>7|7.9</td><td>8|8.3</td><td>5|6.8</td><td>8|7.1</td><td>11|5</td><td>1|4.4</td><td>12|6.2</td><td>40</td></tr>
<tr><td>p-value</td><td>0.870</td><td>0.116</td><td>0.092</td><td>0.259</td><td>0.007</td><td>0.569</td><td><0.001</td><td>0.001</td><td>0.020</td><td></td></tr>
<tr><td>Count</td><td>123</td><td>224</td><td>85</td><td>89</td><td>73</td><td>76</td><td>54</td><td>47</td><td>67</td><td></td></tr>
</table>

With this contingency table containing PMs and problem settings, here defined as a combination of a decision and an assignment level (Sec. 2.3.3), we want to test whether PMs are used in a setting-specific way. Each column represents a separate Fisher test on the null-hypothesis that using a PM or not using a PM is equally likely in each setting. E.g., using patient waiting time in a disc-day problem occurred 7 times, but was expected to be observed 6.3 times (=22x123/429). The p-values in bold represent test results where the null-hypothesis cannot be rejected at a 5% significance level, in which cases we conclude that there is no significant relationship between the setting and the use of a specific PM.

time has been identified as an appropriate PM, it might be crucial to model a rescheduling component. This is the case as rescheduling is primarily used to mitigate overtime. Minimizing overtime in a model that does not include rescheduling does not minimize the real overtime of the hospital, but a function that factors into the hospital's cancellation rate. A more realistic model also includes a rescheduling component that minimizes the cancellation rate.

Generally, it is a problem if the value of the PM is not principally determined by the tested mechanism. In the best case, the model will rightly show that the PM is not influenced by the tested mechanism, which can be a valuable result on its own. Still, including such a PM will shift the focus away from more important PMs, that were not included into the model. In the worst case, the PM will only be dependent on the tested mechanism because of model simplifications (e.g., if deterministic durations are used in open scheduling, surgeon's waiting time is primarily determined by the sequence). In this case, the results derived from the model can suggest benefits that might not be there in reality. Moreover, the implemented mechanisms might worsen the value of other important PMs that were not included in the model.

In order to prevent this problem, the suitability of the PMs to the specific setting should be studied a priory. For example, factor analysis can be used on real data to determine the important factors that determine the value of candidate PMs. It allows to identify the important factors that need to be included in the model in order to get a realistic behavior of the PM (e.g., if the results show that surgery duration uncertainty has a large impact on surgeon waiting time, then optimizing for this PM makes only sense if durations are modeled stochastically). It could also show to what extent the tested mechanism influences the chosen PMs (e.g., whether the sequence of surgeries is amongst the factors that principally determine surgeon's waiting time).

It would be interesting to analyze the connection between problem settings, operations research methods and PMs. This could determine to what extent used PMs are driven by the problem setting (preferred) and to what extent by the method (not preferred), i.e., determine whether PMs are selected because they fit the setting or because they can easily be combined with the chosen method. We tried to uncover these relations using a multiple correspondence analysis. This is a method for decomposing the overall Chi-square statistics, which is similar to decomposing the total variance in Factor Analysis. Unfortunately, this analysis did not yield a result that we could interpret as more than five singular values are needed to only cover 50% of the inertia (in Factor Analysis terms this corresponds to the variance).

6.1.3 Clarifying the problem: Setting- and method-specific assumptions

We found it occasionally difficult to classify some articles as the needed information was either difficult to find or simply not included. Therefore assumptions need to be made clearer. This also allows both researchers and practitioners to more reliably determine whether an article is of interest to them.

Table 6.3 Setting- and method-specific assumptions need to be included in papers on OR planning.

Setting-specific assumptions

Patient characteristics	
Patient type	In/outpatient, emergent, urgent
Duration patterns	Distribution (e.g., log-normal, Empirical), mean/variance
Arrival patterns	Distribution (e.g., Poisson, Empirical), mean/variance
Hospital characteristics	
Capacity size	Nr. ORs, nr. beds, equipment, ...
Capacity pattern	Weekly mean/variance, ...
Personnel	Nr. surgeons, medical staff, ...
Hospital type	General, specialized care, ...
Scheduling policy	Dynamic/static, open/block, ...
Same-day policy	Emergency admittance rules,...
Admission policy	Refusal and deferral policy,...
Problem characteristics	
PM	Waiting time, overtime, leveling, ...
Decision level	Discipline, surgeon, patient, ...
Assignment level	Date, time, room, capacity
Up/downstr. units	ICU, PACU, wards, ...
Planning horizon	4 weeks, 6 months, ...

Method-specific assumptions

Analytical	E.g., Estimated durations are equal (e.g., 1 hour) for all patients, Poisson arrivals, ...
DES	Tested policies represent a good selection of possible policies, ...
MP	E.g., patients to schedule need to be known upfront (static scheduling), surgery cannot start before scheduled start time, ...
Imprv. heur.	E.g., patients to schedule need to be known upfront (static scheduling), ...

We distinguish between setting-specific (often explicit) and method-specific (often implicit) assumptions (Table 6.3).

Setting specific assumptions are key to understand the (extent of) the problem statement. They generally refer to patient, hospital and problem characteristics (Table 6.3). With regards to patients, these include, among others, the distribution of surgery durations or the LOS in downstream units. With regards to the hospital, they mostly cover assumptions on policies and capacity planning, e.g., how many ORs are available for all the surgical disciplines and how are these ORs shared. Finally, with regards to the problem characteristic, they relate, among others, to decisions on whether to incorporate up- or downstream units.

Method-specific assumptions directly result from the chosen operations research method. We find it important to emphasize the necessity to include a description of method-specific assumptions in articles as we noticed that this is not always the case. This is understandable as for researchers who work with a certain methodology for a longer time most of the assumptions are trivial and consequently they are, in comparison to setting-specific assumptions, less consistently reported on. Nevertheless, as they can be difficult to spot by those readers who might have only a limited understanding of the used methodology, we would recommend to highlight them in the text.

There are various assumptions that follow from the chosen method (Table 6.3). An assumption that is typically made when using an MP or an improvement heuristic to solve the patient-to-date assignment problem is that the patient population that needs to be scheduled is known in advance (i.e., at the moment of scheduling). This assumption is often clearly stated and generally also obvious from the problem formulation.

In contrast, there are assumptions that are less obvious and sometimes not clearly mentioned in the text. One such assumption is that surgeries are restricted from starting before their predetermined surgery start time. By including this assumption, the problem formulation of the MP can be simplified. However, this assumption may not always hold in practice as surgeons may start a surgery right after the preceding surgery is finished (e.g., in a setting where one surgeon performs more than one surgery in sequence). In this setting, a method where the next surgery would necessarily need to be kept on hold until its official start time will give wrong results. Consequently, it is important that practitioners are able to clearly identify articles based on this criterion.

An assumption that is often made in analytical methods (e.g., Markov decision processes) is that surgeries correspond to a fixed slot size (e.g., 1 hour). It is important to keep in mind that under this assumption all surgery duration estimates are of equal length. An improved method allows to allocate surgeries of various fixed sizes (e.g., 4 surgeries of 1 hour and 2 of 2 hours) to ORs [112]. It is easy to see that this still results in a very strong assumption as surgery durations are generally estimated on a much finer scale.

Generally, we noticed that method-based assumptions are more difficult to spot for articles where analytical methods, MPs and improvement heuristics are used. In contrast, we found them easier to recognize in articles where DES and constructive algorithms are used as they are methods where a detailed description of the building blocks of the methods is often necessary.

We recommend to clearly mention both setting- and method-specific assumptions in the text (e.g., in a separate section or table). One way of mentioning assumptions in a compact manner is with a classification schema. This idea is already successfully used in queuing theory (Kendall's notation) or machine scheduling (e.g., the three-field notation $\alpha|\beta|\gamma$ [35] describes respectively the machine environment, the job characteristics and the PM). It was first applied to OR scheduling by Cardoen et al. [46]. Future research could focus on expanding this idea.

Next to mentioning the assumptions, they are ideally also motivated both for model simplifications and extensions. If one intends to introduce a model simplification, one must justify that this simplification will not have a major impact on the conclusions (e.g., via data analysis). For example, if one assumes a Poisson distribution for the patient arrival process, one should always show either that the arrival process observed in the hospital is close to a Poisson distribution or that this assumption will only have a small effect on the final conclusions (from our own experience the arrival process can be much more variable than what a Poisson process would suggest).

Similarly, one must also justify extensions to a model (e.g., show via data analysis that the extension represents an important mechanism found in reality). We would recommend to perform this analysis even if the extension was required by the hospital management (i.e., quantify and explore why the management thinks they need those extensions). For example, if one wants to model sequence-dependent turnover times, one has to show that such a dependency can actually be observed in the data of real hospitals and that modeling this dependency in a model brings benefits (e.g., allows to construct more robust schedules).

From the literature, we see that motivations for simplifications or extensions are rare. To include these is important as they ensure that the research remains relevant and that models are kept from getting unnecessarily complex. This is important as simpler models are better. For instance, for a surgery scheduling problem it might be possible to exclude personnel scheduling (e.g., nurse, surgeon, anesthesiology) from the model without any major impact on the validity of the conclusions of the research. However, it could also be the case that excluding the personnel schedule might yield surgery schedules that are infeasible in reality, questioning the validity of the research conclusions.

6.2 Limitations and future work

Our study has several limitations. This is a single center study in a hospital that contains highly utilized ORs. Moreover, the hospital is situated in the Belgian context, which means that waiting times are relatively short [251, 291] and therefore using dynamic scheduling is justified. Moreover, the sequencing step is in our setting not of importance for two reasons. First, surgeons usually own an OR for an entire day and therefore the exact surgery time does not matter (i.e., the next surgery starts when the previous one is finished). Second, due to recent adaptations, the patient flow to downstream facilities is relatively smooth and thus it is unnecessary to sequence patients for an optimized use of downstream facilities. In other settings, in particular at outpatient facilities, the sequence of patients might be highly relevant and should thus be considered.

Many of our results are easy to interpret because scheduling decisions were shown to be independent of OR performance measures (e.g., overtime). This might not necessarily be true in any type of setting. For example, if surgeries cannot start before their planned start time, setting appropriate start times will almost certainly have an effect on OR performance. In those cases a trade-off has to be found between OR- and patient-related performance measures.

Future research can look at whether the conclusions hold in the case that non-electives are served in separate ORs. In those cases, on the one hand, less capacity would remain for electives, while on the other hand, it would bring more stability to the elective schedule. Whether this stability could offset the drawback of having less capacity is of interest to the hospital. We think the same conclusions would hold for this case, as OR-related performance measures would likely remain independent of the used scheduling strategy.

There are some aspects of the real setting that we did not model, but which we also do not deem to be important. They relate, on the one hand, to surgeon and patient preferences and, on the other hand, to downstream facilities. An example of a surgeon preference is to have only one difficult surgery (e.g., hip replacement) on a day. Similarly, also the number of children can be restricted. This is done as patients before their surgery are not allowed to consume food, which is more difficult for children. It is therefore best to serve one child first in the morning. Those factors are important to consider when scheduling patients, but excluding them in the simulation model is unlikely to change the conclusion of our results in a major way.

At some hospitals, capacity problems at downstream facilities such as the ICU and the PACU cause OR blocking and therefore have a detrimental effect on OR usage. It could be interesting to include those aspects into a future version of our simulation model. At the University Hospital Leuven's inpatient OR department blockage at downstream facilities does not pose a problem partly due to recent changes.

In our study we did not include the patient rejection process. Consequently, we only included arrivals that were served by the hospital. Modeling the patients rejection process constitutes an important further extension to our model. This would include patients into the model that originally intended to get surgery at the University Hospital Leuven but ended up getting surgery at another hospital. This new model would consequently use patient data that does not necessarily constitute a feasible schedule, i.e., demand might be higher than supply. Including a patient rejection mechanism into the model can also help us to assess whether the benefits of using the FCFS strategy might be offset by suddenly having more patients requesting surgery.

In the literature, the patient rejection mechanism is generally modeled as a trade-off function where cost factors such as waiting time, OR overtime and OR opening costs are balanced against the profit gained from surgeries. In order to model the patient rejection mechanism of the University Hospital Leuven we need to overcome three challenges.

First, we need to define an appropriate cost structure. This is challenging as besides the academic and medical relevance there are also other considerations determining the value of a surgery. Such values are related to the monetary return (defined by the reimbursement tariff), the fact that a specific expertise is present in a certain hospital and government-related regulations.

Second, it is challenging to model the patient rejection process itself as it is only partly controlled by the hospital. Patients can legally not be rejected. As the University Hospital Leuven is for some surgeries regarded to be the most qualified hospital in the country, patients may be reluctant to go to another hospital even if asked to do so. On the contrary, there might be patients who registered at several hospitals and therefore can cancel at any arbitrary time.

Third, it might be challenging to get a realistic understanding of the rejection process as we are missing the necessary data. Patients that are immediately rejected will not enter the hospital's data system. For example, some patients that went for consultation to a surgeon who has a long waiting list might have been convinced to register for surgery at another hospital. Since they never entered the hospital's data system, we have no information on their number and their attributes. Having this kind of data available would require to convince surgeons to record these patients, which not all of them would be willing to.

A first approach to understand the rejection process at the hospital is to get an understanding of the value surgeons attach to each surgery type. A possible way to get this value is by performing a Delphi study amongst the surgeons.

Chapter 7

Conclusion

In Chapter 2 and Section 6.1 we classified and discussed the OR planning and scheduling literature. We classified the literature with regard to the patient type, the different performance measures, the decisions that have to be made, the integration of up- and downstream units of the OR, the incorporation of uncertainty, the operations research methodology and the testing phase. The resulting classification tables enable the reader to quickly identify new relevant articles (Sec. 2.3.1-2.3.7). Using the classification fields, we found that

- overtime and patient waiting time are the most used performance measures;
- problems on day and room assignments are more often researched than capacity- and timing-related problems;
- although stochastic surgery durations are considered in about 44% of the papers, only 28% of the papers consider stochastic arrivals;
- many authors test the developed approach with real data, but only few report on implementation results in practice;
- a classification matrix, showing both the assignment decisions as well as the decision level (Table 2.5), can help to define the problem characteris-

tics in a less ambiguous way than a terminology-based approach.

We also looked at trends for the last ten years and examined connections between the problem setting, the used methods and the performance measures. This showed that

- the amount of published technical articles has been increasing significantly in the recent ten years;
- surprisingly, research on outpatient surgery is not increasing, despite its increasing importance in reality;
- the amount of papers that investigate the OR in an integrated way (e.g., by including the PACU) is, contrary to what we expected, not increasing;
- MP is the most popular method (included in half of the articles) and its popularity has been increasing over the last ten years;
- analytical and DES models often relate to capacity problems solved at the discipline level. Both generally model the durations and patient arrivals stochastically. Results from DES models, unlike analytical results, are usually tested with real data;
- the number of included performance measures and constraints is the lowest in analytical methods and the highest in DES models;
- most popular constraints are personnel-related (e.g., surgeon availability) and preference-related (e.g., serve higher priority patient first).

We also found that there are no dominant research trends observable. This shows that the research community is not moving into one particular direction, but instead remains occupied with a wide variety of problems and solution methods.

An analysis of the connections between the classification fields showed which methods, PMs and constraints are commonly combined and which are not (Sec. 2.3.8). In general, we see that all combinations have been researched to some extent already. Consequently, one might wonder whether there is anything left in OR planning. One could argue that OR planning is an outdated research topic and the time has come to focus on other research areas. We think that this argument is flawed as the operations research community did not fulfill its job yet.

In particular, we suggest two directions for future research. First, there are still new topics to be further explored (e.g., bulking behavior of patients, case studies on the effect of modeling assumptions, integration of the inpatient and outpatient schedules). Second, there are popular problems for which the proposed solutions have not yet been adopted by the stakeholders and therefore need to be revisited.

In Section 6.1, we identified ways to get results that are more applied by stakeholders. We therefore identified common pitfalls and points that, based on our analysis of the literature, deserve special attention when researching this field. We found that

- there is a need for a clearer distinction between theoretic articles that contribute advanced methods and applied articles that show the real-life applicability of these methods (Table 6.1). This distinction would allow articles to focus on their core strengths. Additionally, it would make it easier for both practitioners and researchers to identify articles that are relevant for them;
- many PMs (e.g., overtime) are used in articles indifferently of the tackled problem. As a consequence, the most appropriate PMs for a setting are not necessarily the ones that are the most popular in the literature;
- important information is occasionally missing from articles. This makes it harder for readers (especially practitioners) to determine whether the shown research results are relevant to them. For example, generally articles where analytical methods (e.g., Markov models) are used, will often assume estimated durations to be equal. As this is a strong assumption, one should be careful when generalizing the results of these methods to inpatient scheduling.

In order to avoid these pitfalls, we conclude that researchers need to

- decide on whether researchers or practitioners are targeted (Sec. 6.1.1) and, based on this decision, follow the respective guidelines from Table 6.1. The guidelines make important differences between the target groups with regard to the defined problem setting, the research task, the reported findings and conclusions. The choice of targeting researchers or practitioners should be clear and consistent throughout the article. This distinction also requires adapted publishing incentives;

- choose adequate PMs keeping in mind two steps (Sec. 6.1.2). First, choose PMs that are of practical relevance to the stakeholders of the hospital. Second, check whether the model components that principally drive the chosen PMs are modeled. For example, overtime might not be determined by the patient-to-date assignment policy, but rather by how well surgeons estimated their surgery durations. In this case it is important to include a realistic surgery duration model;

- mention (and ideally motivate) the setting- and method-specific assumptions, as outlined in Table 6.3, clearly in the article (Sec. 6.1.3). Clarifying method-specific assumptions is particularly important since the readers might not always be familiar with the used operations research methods. Spelling out all assumptions helps them to understand whether a method or result is of relevance to them.

In Chapters 3, 4 and 5 and Sections 3.2 and 6.2 we discussed the OR scheduling setting of the University Hospital Leuven and showed the results of applying various surgery scheduling methods. We determined that the simulation model of the OR department needs to include non-electives since they have a large impact on the OR department and the elective schedule. This is the case as non-electives instead of entering an arbitrary empty or low utilized OR will often be assigned to ORs that accommodate the corresponding discipline (Fig. 3.2). These ORs can already be fully planned with electives. This combined with the fact that some disciplines need to serve a large amount of non-electives (Fig. 3.5) means that their ORs can be heavily utilized. This can lead to overtime and elective rescheduling.

Elective rescheduling is a component that also needs to be included into the simulation model. It contributes to the fact that the overtime, the undertime and the utilization of ORs will not depend on the chosen patient scheduling strategy as often assumed in the literature. The rescheduling model determines, firstly, how patients can be reassigned to ORs of different disciplines and, secondly, imitates the timing of decisions made in reality. We found that in reality surgery reassignments and cancellations happen continuously throughout the day.

We also found that in order to realistically model surgery durations in an inpatient setting the following rules are important. If only realized durations are modeled on the pathology level, the log-logistic distribution should be used. If realized durations are modeled on an aggregated level (e.g., discipline), we advice to use a GMM. If both realized and estimated durations are modeled, then either a fully empirical distribution should be chosen or a bivariate copula model that is able to handle multimodality (e.g., GMC). The marginal distribution of the copula model should be based on a GMM or a KDE. If estimated durations contain a pronounced discrete component, the corresponding marginal distribution should be based on a KDE.

There are straightforward managerial implications of our results for the surgeons of the hospital and for schedulers of similar hospitals. In order to serve a large number of patients within their DT it is more important to focus on the efficient use of OR capacities rather than on patient priorities. Therefore, FCFS, which is a strategy that makes good use of OR capacities, will perform well. FCFS might not always be applicable in reality as patients from less urgent urgency classes may not always be available for surgery on short notice. Therefore, it is important to allow for patient replanning. One might intuitively think that replanning high priority patients is advantageous, but we showed that it is better to replan those ones that best fill out free next day OR capacity. This is the case as the major benefit of replanning stems from saving OR capacity, i.e., valuable capacity from the originally planned date is exchanged for less valuable next day capacity that is in danger of being wasted. Efficient use of capacity also entails that surgeons should in a timely manner release ORs that they do not use so that other surgeons can use them.

We also tested a two-step strategy where surgeries are assigned to a week first, and only in a second step to an OR and a weekday. We have shown that protection levels, a strategy that reserves capacity for high urgency patients, preforms not as expected since it does not decrease the percentage of patients served within their DT, increases patient waiting time and decreases the weighted DT cost only if combined with the push factor. We have also shown that it is both beneficial to consider the DT during the within-week scheduling step and to allow patients to be pushed into the fixed weekly schedule.

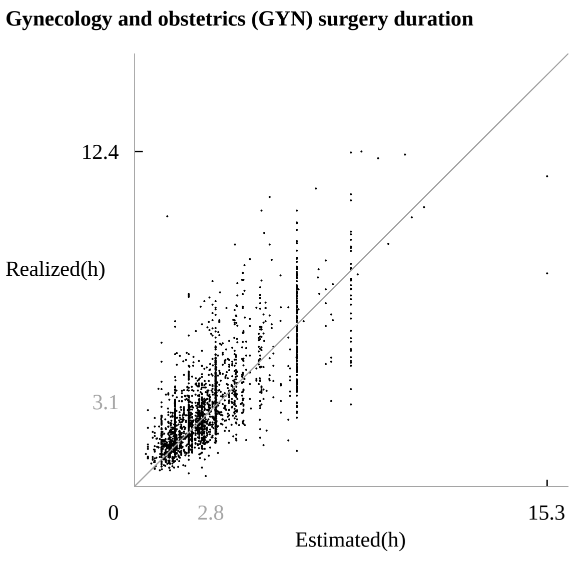
Our results also entail that evaluating the capacity assigned to surgeons should not only be based on the average patient waiting time but also on the used scheduling strategy. This is the case as low average waiting times do not necessarily mean that surgeons have more capacity available than needed, but it can also mean that they use a scheduling strategy where patients are scheduled to a date close to their arrival date and therefore capacity is used efficiently. This behavior should not be punished. Similarly, Dexter et al. [73] argue that determining an appropriate amount of block time and selecting a method to schedule cases into a surgeon's blocks must be done simultaneously.

Results that are specific to the case hospital, show that DT category 4 patients are in reality overprioritized. Giving more importance to other DT categories when creating the surgery schedule could lead to larger gains with regards to the weighted DT cost. We think that following these and previously made recommendations can lead to better schedules and therefore further increase the quality of care provided to patients by the hospital.

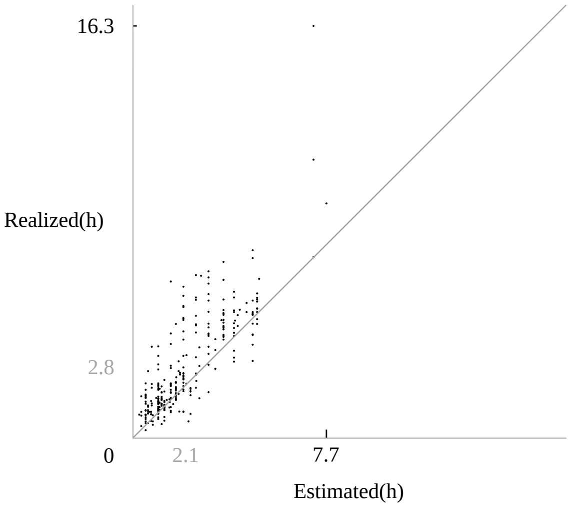
Appendices

Appendix A

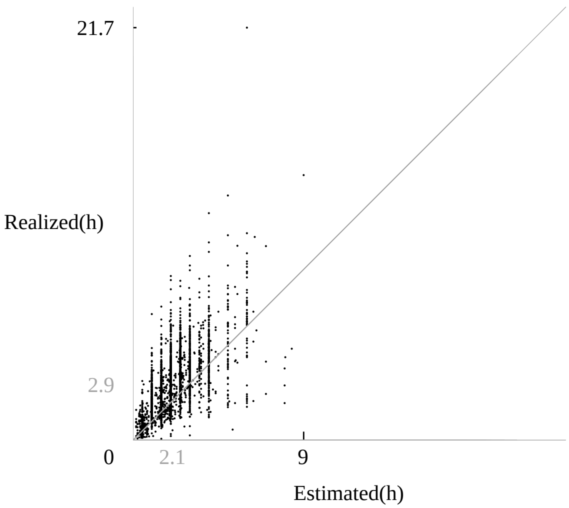
Surgery durations



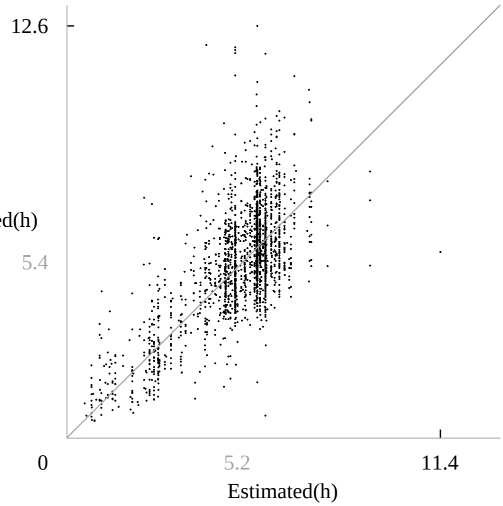
Abdominal transplant (Tx) surgery duration



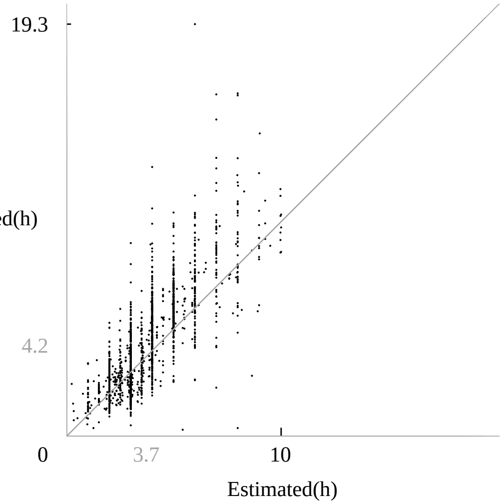
Abdominal (ABD) surgery duration



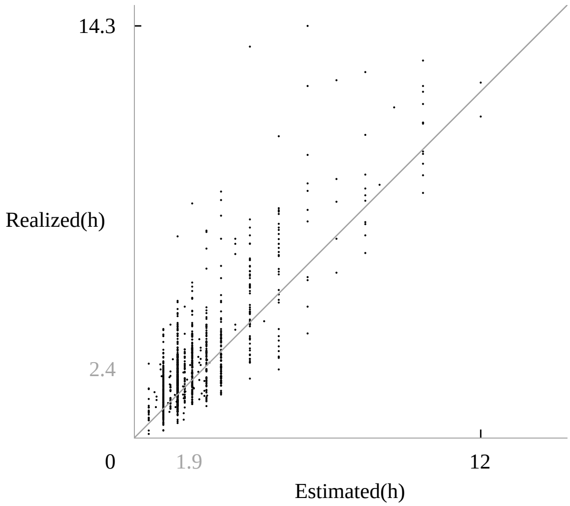
Cardiac (CAH) surgery duration



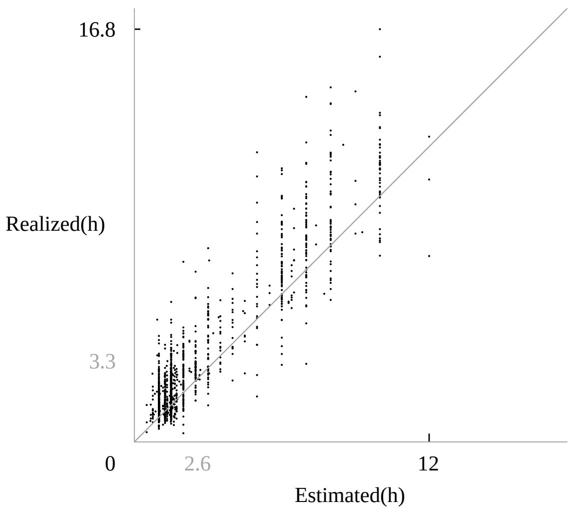
Neur (NCH) surgery duration



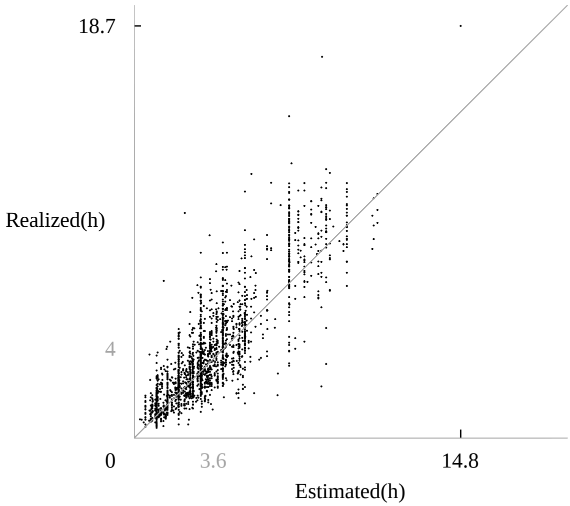
General medical oncological (ONC) surgery duration



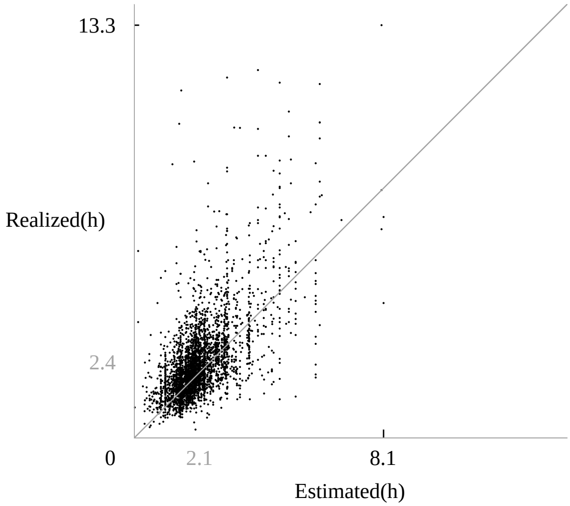
Plastic, reconstructive and cosmetic (RHK) surgery duration



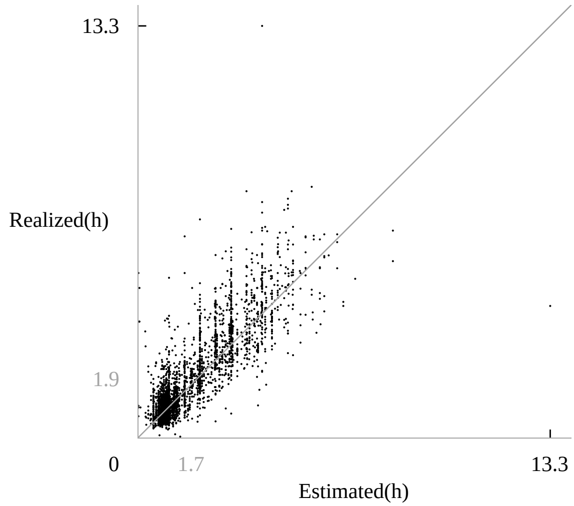
Thoracic (THO) surgery duration



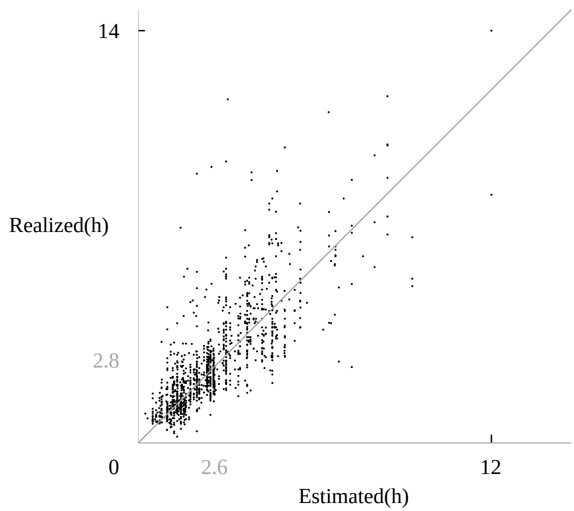
Traumatology (TRH) surgery duration



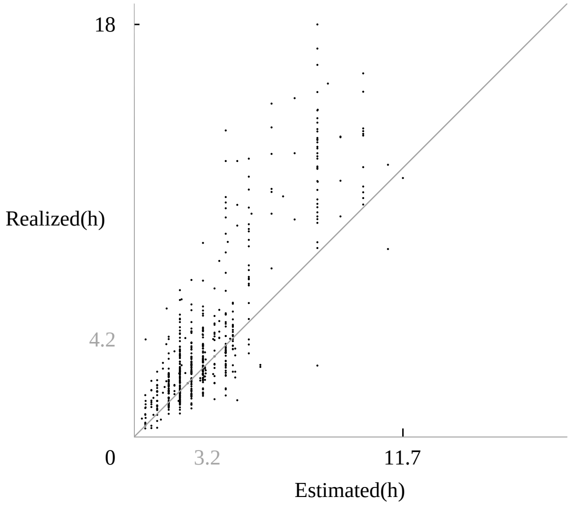
Urology (URO) surgery duration



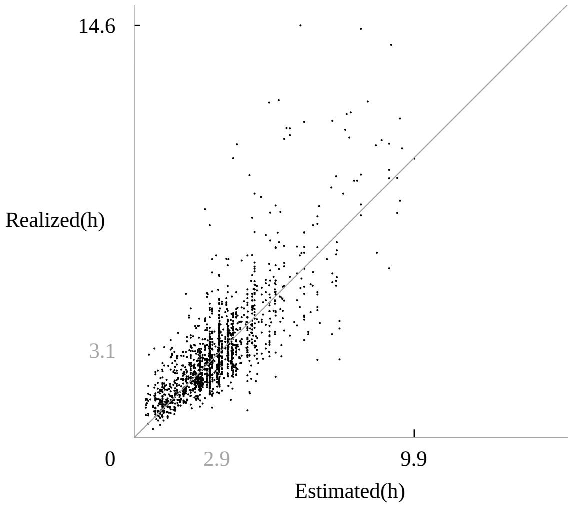
Vascular (VAT) surgery duration



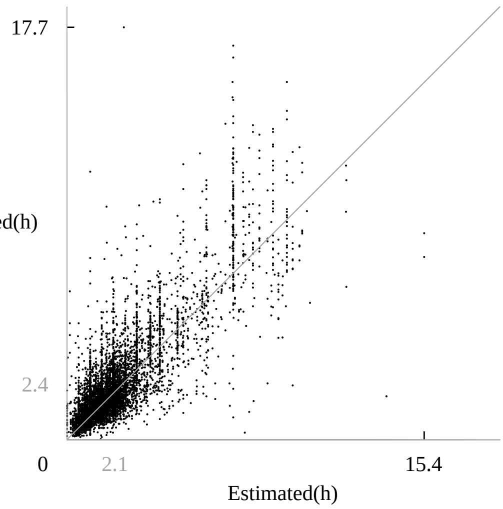
Oral and maxillofacial (MKA) surgery duration



Head and neck (NKO) surgery duration



Non-elective (EMG) surgery duration



Appendix B

Surgeon estimation error

Table B.1 Is the systematic underestimation of surgery durations a consequence of many surgeons underestimating by a little or is it a consequence of a few surgeons underestimating by a lot? The answer to this question depends on the discipline. For most disciplines, many surgeons underestimate by a little. However, there are exceptions, for MKA one surgeon is causing 78% of total underestimated hours and is underestimating surgeries by 31%. Also ONC, TRH, and URO, have one one surgeon who causes a particularly large amount of underestimated surgery duration hours while also substantially underestimating surgeries.

	The percentage of the total underestimated hours caused by a surgeon (to the left of 'ǀ') and the percentage the surgeon underestimates surgeries in average (to the right of 'ǀ')				
	Surgeon 1	Surgeon 2	Surgeon 3	#Surgeons	#Surgeons >20 yearly surgeries
GYN	25% ǀ 13%	23% ǀ 18%	16% ǀ 12%	15	9
Tx	52% ǀ 27%	47% ǀ 23%	1% ǀ 15%	3	2
ABD	35% ǀ 32%	23% ǀ 31%	15% ǀ 37%	8	8
CAH	35% ǀ 7%	33% ǀ 4%	18% ǀ 2%	6	5
NCH	28% ǀ 19%	26% ǀ 22%	24% ǀ 19%	8	7
ONC	**67% ǀ 25%**	29% ǀ 20%	3% ǀ 52%	7	3
RHK	21% ǀ 27%	20% ǀ 24%	18% ǀ 26%	6	6
THO	30% ǀ 11%	23% ǀ 17%	22% ǀ 10%	6	6
TRH	**45% ǀ 15%**	25% ǀ 13%	20% ǀ 15%	5	5
URO	**40% ǀ 18%**	22% ǀ 11%	20% ǀ 15%	10	8
VAT	45% ǀ 7%	43% ǀ 8%	13% ǀ 12%	4	4
MKA	**78% ǀ 31%**	19% ǀ 24%	3% ǀ 24%	5	3
NKO	46% ǀ 8%	28% ǀ 13%	12% ǀ 6%	9	6
EMG	10% ǀ 20%	8% ǀ 16%	6% ǀ 18%	90	42

We only show the three surgeons that caused the largest percentage of the total underestimated hours. Surgeons who cause at least 40% of the total underestimated hours, underestimate surgeries by over 10% and belong to a discipline with at least 5 surgeons are shown in bolt.

List of figures

List of tables

Bibliography

[1] Abdelrasol Z, Harraz N, Eltawil A (2014) Operating room scheduling problems: A survey and a proposed solution framework, Springer Netherlands

[2] Adan I, Bekkers J, Dellaert N, Vissers J, Yu XT (2009) Patient mix optimisation and stochastic resource requirements: A case study in cardiothoracic surgery planning. Health Care Management Science 12:129–141

[3] Adan I, Bekkers J, Dellaert N, Jeunet J, Vissers J (2011) Improving operational effectiveness of tactical master plans for emergency and elective patients under stochastic demand and capacitated resources. European Journal of Operational Research 213:290–308

[4] Agnetis A, Coppi A, Corsini M, Dellino G, Meloni C, Pranzo M (2012) Long term evaluation of operating theater planning policies. Operations Research for Health Care 1:95–104

[5] Agnetis A, Coppi A, Corsini M, Dellino G, Meloni C, Pranzo M (2014) A decomposition approach for the combined master surgical schedule and surgical case assignment problems. Health Care Management Science 17:49–59

[6] AIHW (2013) Australian hospital statistics: National emergency access and elective surgery targets 2012. Tech. rep., Australian Institute of Health and Welfare

[7] AIHW (2013) Australian hospital statistics: National emergency access and elective surgery targets 2012. Tech. rep., Australian Institute of Health and Welfare

[8] Al-Amin M, Housman M (2012) Ambulatory surgery center and general hospital competition: Entry decisions and strategic choices. Health Care Management Review 37:223–234

[9] Antonelli D, Taurino T (2010) Application of a patient flow model to a surgery department. In: 2010 IEEE Workshop on Health Care Management (WHCM)

[10] Argo JL, Vick CC, Graham LA, Itani KMF, Bishop MJ, Hawn MT (2009) Elective surgical case cancellation in the veterans health administration system: Identifying areas for improvement. American Journal of Surgery 198:600–606

[11] Arnaout JPM, Kulbashian S (2008) Maximizing the utilization of operating rooms with stochastic times using simulation. In: Proceedings of the 2008 Winter Simulation Conference, pp 1617–1623

[12] Augusto V, Xie X, Perdomo V (2008) Operating theatre scheduling using lagrangian relaxation. European Journal of Industrial Engineering 2:172–189

[13] Augusto V, Xie X, Perdomo V (2010) Operating theatre scheduling with patient recovery in both operating rooms and recovery beds. Computers & Industrial Engineering 58:231–238

[14] Azari-Rad S, Yontef AL, Aleman DM, Urbach DR (2013) Reducing elective general surgery cancellations at a Canadian hospital. Canadian Journal of Surgery 56:113–118

[15] Ballard SM, Kuhl ME (2006) The use of simulation to determine maximum capacity in the surgical suite operating room. In: Proceedings of the 2006 Winter Simulation Conference, pp 433–438

[16] Banditori C, Cappanera P, Visintin F (2013) A combined optimization-simulation approach to the master surgical scheduling problem. IMA Journal of Management Mathematics 24:155–187

[17] Banditori C, Cappanera P, Visintin F (2014) Investigating the relationship between resources balancing and robustness in master surgical scheduling. In: Proceedings of the International Conference on Health Care Systems Engineering, vol 61, pp 149–162

[18] Barkaoui K, Dechambre P, Hachicha R (2002) Verification and optimisation of an operating room workflow. Proceedings of the 35th Annual Hawaii International Conference on System Sciences pp 2581–90

[19] Barua B, Esmail N (2013) Waiting your turn: Wait times for health care in Canada. Tech. rep., Fraser Institute

[20] Barua B, Esmail N (2013) Waiting your turn: Wait times for health care in Canada. Tech. rep., Fraser Institute

[21] Basson MD, Butler T (2006) Evaluation of operating room suite efficiency in the veterans health administration system by using data-envelopment analysis. American Journal of Surgery 192:649–656

[22] Batun S, Denton BT, Huschka TR, Schaefer AJ (2011) Operating room pooling and parallel surgery processing under uncertainty. INFORMS Journal on Computing 23:220–237

[23] Baumgart A, Zoeller A, Denz C, Bender HJ, Heinzl A, Badreddin E (2007) Using computer simulation in operating room management: Impacts on process engineering and performance. In: Proceedings of the 40th Annual Hawaii International Conference on System Sciences, p 10

[24] Beliën J, Demeulemeester E (2007) Building cyclic master surgery schedules with leveled resulting bed occupancy. European Journal of Operational Research 176:1185–1204

[25] Beliën J, Demeulemeester E (2007) Building cyclic master surgery schedules with leveled resulting bed occupancy. European Journal of Operational Research 176:1185–1204

[26] Beliën J, Demeulemeester E (2008) A branch-and-price approach for integrating nurse and surgery scheduling. European Journal of Operational Research 189:652–668

[27] Beliën J, Demeulemeester E, Cardoen B (2006) Visualizing the demand for various resources as a function of the master surgery schedule: A case study. Journal of Medical Systems 30:343–50

[28] Beliën J, Demeulemeester E, Cardoen B (2009) A decision support system for cyclic master surgery scheduling with multiple objectives. Journal of Scheduling 12:147–161

[29] Berg B, Denton BT, Erdogan SA, Rohleder T, Huschka TR (2014) Optimal booking and scheduling in outpatient procedure centers. Computers & Operations Research 50:24–37

[30] Blake JT (2011) Capacity planning in operating rooms, CRG Press, pp 34.1–34.12

[31] Blake JT, Carter MW (1997) Surgical process scheduling: A structured review. Journal of Society for Health Systems 5:17–30

[32] Blake JT, Carter MW (2002) A goal programming approach to strategic resource allocation in acute care hospitals. European Journal of Operational Research 140:541–561

[33] Blake JT, Donald J (2002) Mount Sinai hospital uses integer programming to allocate operating room time. Interfaces 32:63–73

[34] Blake JT, Dexter F, Donald J (2002) Operating room managers' use of integer programming for assigning block time to surgical groups: A case study. Anesthesia and Analgesia 94:143–148

[35] Blazewicz J, Lenstra JK, Rinnooy Kan AHG (1983) Scheduling subject to resource constraints: Classification and complexity. Discrete Applied Mathematics 5:11–24

[36] Boldy D (1976) Review of application of mathematical-programming to tactical and strategic health and social-services problems. Operational Research Quarterly 27:439–448

[37] Bowers J (2013) Balancing operating theatre and bed capacity in a cardiothoracic centre. Health Care Management Science 16:236–244

[38] Bowers J, Mould G (2004) Managing uncertainty in orthopaedic trauma theatres. European Journal of Operational Research 154:599–608

[39] Bowers J, Mould G (2005) Ambulatory care and orthopaedic capacity planning. Health Care Management Science 8:41–7

[40] Brailsford SC, Vissers J (2011) OR in healthcare: A European perspective. European Journal of Operational Research 212:223–234

[41] Brailsford SC, Bolt TB, Bucci G, Chaussalet TM, Connell NA, Harper PR, Klein JH, Pitt M, Taylor M (2013) Overcoming the barriers: A qualitative study of simulation adoption in the NHS. Journal of the Operational Research Society 64:157–168

[42] Bruni ME, Beraldi P, Conforti D (2014) A stochastic programming approach for operating theatre scheduling under uncertainty. IMA Journal of Management Mathematics

[43] Calichman MV (2005) Creating an optimal operating room schedule. AORN Journal 81:580–8

[44] Cappanera P, Visintin F, Banditori C (2014) Comparing resource balancing criteria in master surgical scheduling: A combined optimisation-simulation approach. International Journal of Production Economics 158:179–196

[45] Cardoen B, Demeulemeester E (2008) Capacity of clinical pathways: A strategic multi-level evaluation tool. Journal of Medical Systems 32:443–452

[46] Cardoen B, Demeulemeester E (2010) Operating room planning and scheduling: A classification scheme. International Journal of Health Management and Information 1:71–83

[47] Cardoen B, Demeulemeester E (2011) A decision support system for surgery sequencing at UZ leuven's day-care department. International Journal of Information Technology & Decision Making 10:435–450

[48] Cardoen B, Demeulemeester E, Beliën J (2009) Optimizing a multiple objective surgical case sequencing problem. International Journal of Production Economics 119:354–366

[49] Cardoen B, Demeulemeester E, Beliën J (2009) Sequencing surgical cases in a day-care environment: An exact branch-and-price approach. Computers & Operations Research 36:2660–2669

[50] Cardoen B, Demeulemeester E, Beliën J (2010) Operating room planning and scheduling: A literature review. European Journal of Operational Research 201:921–932

[51] Cardoen B, Demeulemeester E, Van der Hoeven J (2010) On the use of planning models in the operating theatre: Results of a survey in Flanders. International Journal of Health Planning and Management 25:400–414

[52] Cardoen B, Demeulemeester E, Van der Hoeven J (2010) On the use of planning models in the operating theatre: Results of a survey in Flanders. International Journal of Health Planning and Management 25:400–414

[53] Carter MW, Ketabi S (2013) Bed balancing in surgical wards via block scheduling. Journal of Minimally Invasive Surgical Sciences 2:129–137

[54] Cayirli T, Veral E (2003) Outpatient scheduling in health care: A review of literature. Production and Operations Management 12:519–549

[55] Ceschia S, Schaerf A (2014) Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays. Journal of Scheduling pp 1–13

[56] Chaabane S, Meskens N, Guinet A, Laurent M (2006) Comparison of two methods of operating theatre planning: Application in Belgian hospital. In: 2006 International Conference on Service Systems and Service Management, pp 386–392

[57] Chang JH, Chen KW, Chen KB, Poon KS, Liu SK (2014) Case review analysis of operating room decisions to cancel surgery. BMC Surgery 14:47

[58] Choi S, Wilhelm WE (2014) An approach to optimize block surgical schedules. European Journal of Operational Research 235:138–148

[59] Choi S, Wilhelm WE (2014) On capacity allocation for operating rooms. Computers & Operations Research 44:174–184

[60] Conforti D, Guerriero F, Guido R (2010) A multi-objective block scheduling model for the management of surgical operating rooms: New solution approaches via genetic algorithms. In: 2010 IEEE Workshop on Health Care Management (WHCM), p 5 pp.

[61] Creemers S, Beliën J, Lambrecht M (2012) The optimal allocation of server time slots over different classes of patients. European Journal of Operational Research 219:508–521

[62] Day R, Garfinkel R, Thompson S (2012) Integrated block sharing: A win-win strategy for hospitals and surgeons. M&Som-Manufacturing & Service Operations Management 14:567–583

[63] Dekhici L, Belkadi K (2010) Operating theatre scheduling under constraints. Journal of Applied Sciences pp 1380–8

[64] Dellaert N, Jeunet J (2008) Hospital admission planning to optimize major resources utilization under uncertainty. In: 3rd World Conference on Production and Operations Management, p 16

[65] Demeulemeester E, Beliën J, Cardoen B, Samudra M (2013) Operating room planning and scheduling, Springer

[66] Denton BT, Gupta D (2003) A sequential bounding approach for optimal appointment scheduling. IIE Transactions 35:1003–1016

[67] Denton BT, Rahman AS, Nelson H, Bailey AC (2006) Simulation of a multiple operating room surgical suite. In: Proceedings of the 2006 Winter Simulation Conference, pp 414–424

[68] Denton BT, Viapiano J, Vogl A (2007) Optimization of surgery sequencing and scheduling decisions under uncertainty. Health Care Management Science 10:13–24

[69] Denton BT, Miller AJ, Balasubramanian HJ, Huschka TR (2010) Optimal allocation of surgery blocks to operating rooms under uncertainty. Operations Research 58:802–816

[70] Dexter F (2000) A strategy to decide whether to move the last case of the day in an operating room to another empty operating room to decrease overtime labor costs. Anesthesia and Analgesia 91:925–928

[71] Dexter F, Epstein RH (2009) Typical savings from each minute reduction in tardy first case of the day starts. Anesthesia and Analgesia 108:1262–1267

[72] Dexter F, Traub RD (2002) How to schedule elective surgical cases into specific operating rooms to maximize the efficiency of use of operating room time. Anesthesia and Analgesia 94:933–942

[73] Dexter F, Macario A, Traub RD, Hopwood M, Lubarsky DA (1999) An operating room scheduling strategy to maximize the use of operating room block time: computer simulation of patient scheduling and survey of patients' preferences for surgical waiting time. Anesth Analg 89:7–20

[74] Dexter F, Macario A, O'Neill L (2000) Scheduling surgical cases into overflow block time: Computer simulation of the effects of scheduling strategies on operating room labor costs. Anesthesia and Analgesia 90:980–988

[75] Dexter F, Macario A, Lubarsky DA (2001) The impact on revenue of increasing patient volume at surgical suites with relatively high operating room utilization. Anesthesia and Analgesia 92:1215–1221

[76] Dexter F, Blake JT, Penning DH, Sloan B, Chung P, Lubarsky DA (2002) Use of linear programming to estimate impact of changes in a hospital's operating room time allocation on perioperative variable costs. Anesthesiology 96:718–724

[77] Dexter F, Lubarsky DA, Blake JT (2002) Sampling error can significantly affect measured hospital financial performance of surgeons and resulting operating room time allocations. Anesthesia and Analgesia 95:184–188

[78] Dexter F, Traub RD, Macario A (2003) How to release allocated operating room time to increase efficiency: Predicting which surgical service will have the most underutilized operating room time. Anesthesia and Analgesia 96:507–512

[79] Dexter F, Epstein RH, Traub RD, Xiao Y (2004) Making management decisions on the day of surgery based on operating room efficiency and patient waiting times. Anesthesiology 101:1444–1453

[80] Dexter F, Ledolter J, Wachtel RE (2005) Tactical decision making for selective expansion of operating room resources incorporating financial criteria and uncertainty in subspecialties' future workloads. Anesthesia and Analgesia 100:1425–1432

[81] Dexter F, Macario A, Ledolter J (2007) Identification of systematic underestimation (bias) of case durations during case scheduling would not markedly reduce overutilized operating room time. Journal of Clinical Anesthesia 19:198–203

[82] Dexter F, Birchansky L, Bernstein JM, Wachtel RE (2009) Case scheduling preferences of one surgeon's cataract surgery patients. Anesthesia and Analgesia 108:579–582

[83] Dexter F, Wachtel RE, Epstein RH, Ledolter J, Todd MM (2010) Analysis of operating room allocations to optimize scheduling of specialty rotations for anesthesia trainees. Anesthesia and Analgesia 111:520–524

[84] Dexter F, Maxbauer T, Stout C, Archbold L, Epstein RH (2014) Relative influence on total cancelled operating room time from patients who are inpatients or outpatients preoperatively. Anesth Analg 118:1072–80

[85] Di Martinelly C, Baptiste P, Maknoon MY (2014) An assessment of the integration of nurse timetable changes with operating room planning and scheduling. International Journal of Production Research 52:7239–7250

[86] Does R, Vermaat TMB, Verver JPS, Bisgaard S, Van den Heuvel J (2009) Reducing start time delays in operating rooms. Journal of Quality Technology 41:95–109

[87] Epstein RH, Dexter F (2013) Rescheduling of previously cancelled surgical cases does not increase variability in operating room workload when cases are scheduled based on maximizing efficiency of use of operating room time. Anesth Analg 117:995–1002

[88] Epstein RH, Dexter F (2015) Management implications for the perioperative surgical home related to inpatient case cancellations and add-on case scheduling on the day of surgery. Anesthesia and Analgesia 121:206–18

[89] Epstein RH, Dexter F (2015) Management implications for the perioperative surgical home related to inpatient case cancellations and add-on case scheduling on the day of surgery. Anesth Analg 121:206–18

[90] Erdem E, Qu X, Shi J (2012) Rescheduling of elective patients upon the arrival of emergency patients. Decision Support Systems 54:551–563

[91] Erdogan SA, Denton BT (2010) Surgery planning and scheduling, John Wiley & Sons, Inc.

[92] van Essen JT, Hans EW, Hurink JL, Oversberg A (2012) Minimizing the waiting time for emergency surgery. Operations Research for Health Care 1:34–44

[93] van Essen JT, Hurink JL, Hartholt W, van den Akker BJ (2012) Decision support system for the operating room rescheduling problem. Health Care Management Science 15:355–372

[94] van Essen JT, Bosch JM, Hans EW, Van Houdenhoven M, Hurink JL (2014) Reducing the number of required beds by rearranging the OR-schedule. OR Spectrum 36:585–605

[95] Everett JE (2002) A decision support simulation model for the management of an elective surgery waiting system. Health Care Management Science 5:89–95

[96] Ewen H, Mönch L (2014) A simulation-based framework to schedule surgeries in an eye hospital. IIE Transactions on Healthcare Systems Engineering 4:191–208

[97] Ewing J (2014) Revenue growth and cash flow margins hit all-time lows in 2013 US not-for-profit hospital medians. Tech. rep., Moody's

[98] Fei H, Meskens N, Chu C (2007) An operating theatre planning and scheduling problem in the case of a "block scheduling" strategy. In: 2006 International Conference on Service Systems and Service Management

[99] Fei H, Chu C, Meskens N, Artiba A (2008) Solving surgical cases assignment problem by a branch-and-price approach. International Journal of Production Economics 112:96–108

[100] Fei H, Chu C, Meskens N (2009) Solving a tactical operating room planning problem by a column-generation-based heuristic procedure with four criteria. Annals of Operations Research 166:91–108

[101] Fei H, Meskens N, Combes C, Chu C (2009) The endoscopy scheduling problem: A case study with two specialised operating rooms. International Journal of Production Economics 120:452–462

[102] Fei H, Meskens N, Chu C (2010) A planning and scheduling problem for an operating theatre using an open scheduling strategy. Computers & Industrial Engineering 58:221–230

[103] Fei H, Meskens N, El-Darzi E (2010) Evaluating alternative surgery plans with discrete-event simulation model. In: 2010 IEEE Workshop on Health Care Management (WHCM), pp 1–6

[104] Ferrand Y, Magazine M, Rao U (2010) Comparing two operating-room-allocation policies for elective and emergency surgeries. In: Proceedings of the 2010 Winter Simulation Conference, pp 2364–74

[105] Ferrand Y, Magazine M, Rao U (2014) Partially flexible operating rooms for elective and emergency surgeries. Decision Sciences 45:819–847

[106] Ferrand YB, Magazine MJ, Rao US (2014) Managing operating room efficiency and responsiveness for emergency and elective surgeries: A literature survey. IIE Transactions on Healthcare Systems Engineering 4:49–64

[107] Ferrin DM, Miller MJ, Wininger S, Neuendorf MS (2004) Analyzing incentives and scheduling in a major metropolitan hospital operating room through simulation. In: Proceedings of the 2004 Winter Simulation Conference, pp 1975–1980

[108] Fügener A, Hans EW, Kolisch R, Kortbeek N, Vanberkel PT (2014) Master surgery scheduling with consideration of multiple downstream units. European Journal of Operational Research 239:227–236

[109] Fischetti M, Monaci M (2009) Light Robustness, vol 5868, Springer Berlin Heidelberg, pp 61–84

[110] Gartner D, Kolisch R (2014) Scheduling the hospital-wide flow of elective patients. European Journal of Operational Research 233:689–699

[111] Ghazalbash S, Sepehri MM, Shadpour P, Atighehchian A (2012) Operating room scheduling in teaching hospitals. Advances in Operations Research 2012:16

[112] Gocgun Y, Ghate A (2012) Lagrangian relaxation and constraint generation for allocation and advanced scheduling. Computers & Operations Research 39:2323–2336

[113] Gomes C, Almada-Lobo B, Borges J, Soares C (2012) Integrating data mining and optimization techniques on surgery scheduling, vol 7713, Springer Berlin Heidelberg, pp 589–602

[114] Gonzalez P, Herrero C (2004) Optimal sharing of surgical costs in the presence of queues. Mathematical Methods of Operations Research 59:435–446

[115] Guda H, Dawande M, Janakiraman G, Jung KS (2016) Optimal policy for a stochastic scheduling problem with applications to surgical scheduling. Production and Operations Management

[116] Guerriero F, Guido R (2011) Operational research in the management of the operating theatre: A survey. Health Care Management Science 14:89–114

[117] Guinet A, Chaabane S (2003) Operating theatre planning. International Journal of Production Economics 85:69–81

[118] Gul S, Denton BT, Fowler JW, Huschka TR (2011) Bi-criteria scheduling of surgical services for an outpatient procedure center. Production and Operations Management 20:406–417

[119] Gul S, Denton B, Fowler JW (2012) A multi-stage stochastic integer programming model for surgery planning, Michigan Engineering

[120] Gupta D (2007) Surgical suites' operations management. Production and Operations Management 16:689–700

[121] Gupta D, Denton BT (2008) Appointment scheduling in health care: Challenges and opportunities. IIE Transactions 40:800–819

[122] Gupta D, Natarajan MK, Gafni A, Wang L, Shilton D, Holder D, Yusuf S (2007) Capacity planning for cardiac catheterization: A case study. Health Policy 82:1–11

[123] Hans EW, Vanberkel PT (2012) Operating theatre planning and scheduling, vol 168, Springer US, pp 105–130

[124] Hans EW, Nieberg T, van Oostrum JM (2007) Optimization in surgery planning. Medium Econometrische Toepassingen 15

[125] Hans EW, Wullink G, Van Houdenhoven M, Kazemier G (2008) Robust surgery loading. European Journal of Operational Research 185:1038–1050

[126] Hans EW, Wullink G, Van Houdenhoven M, Kazemier G (2008) Robust surgery loading. European Journal of Operational Research 185:1038–1050

[127] Hanset A, Meskens N, Duvivier D (2010) Using constraint programming to schedule an operating theatre. In: 2010 IEEE Workshop on Health Care Management (WHCM)

[128] Harper PR (2002) A framework for operational modelling of hospital resources. Health Care Management Science 5:165–173

[129] Hashemi D, Seyed H, Rousseau LM, Pesant G (2014) A constraint programming-based column generation approach for operating room planning and scheduling, vol 8451, Springer International Publishing, pp 455–463

[130] Heng M, Wright JG (2013) Dedicated operating room for emergency surgery improves access and efficiency. Canadian Journal of Surgery 56:167–174

[131] Herring WL, Herrmann JW (2012) The single-day surgery scheduling problem: Sequential decision-making and threshold-based heuristics. OR Spectrum 34:429–459

[132] HFMA (2003) Achieving operating room efficiency through process integration. Tech. rep., Healthcare Financial Management Association

[133] HFMA (2011) Value in health care: Current state and future directions. Tech. rep., Healthcare Financial Management Association

[134] Holte M, Mannino C (2013) The implementor/adversary algorithm for the cyclic and robust scheduling problem in health-care. European Journal of Operational Research 226:551–559

[135] Hosseini N, Taaffe KM (2014) Allocating operating room block time using historical caseload variability. Health Care Management Science

[136] Hosseini N, Taaffe KM (2014) Allocating operating room block time using historical caseload variability. Health Care Management Science

[137] Hsu VN, de Matta R, Lee CY (2003) Scheduling patients in an ambulatory surgical center. Naval Research Logistics 50:218–238

[138] Huh WT, Liu N, Van-Anh T (2013) Multiresource allocation scheduling in dynamic environments. M&Som-Manufacturing & Service Operations Management 15:280–291

[139] Hulshof P, Boucherie RJ, van Essen JT, Hans EW, Hurink JL, Kortbeek N, Litvak N, Vanberkel PT, van der Veen E, Veltman B, Vliegen IMH, Zonderland ME (2011) ORchestra: an online reference database of OR/MS literature in health care. Health Care Management Science 14:383–384

[140] Hulshof P, Kortbeek N, Boucherie RJ, Hans EW, Bakker P (2012) Taxonomic classification of planning decisions in health care: A structured review of the state of the art in OR/MS. Health Systems 1:129–175

[141] Hulshof P, Boucherie RJ, Hans EW, Hurink JL (2013) Tactical resource allocation and elective patient admission planning in care processes. Health Care Management Science 16:152–66

[142] Huschka TR, Denton BT, Gul S, Fowler JW (2007) Bi-criteria evaluation of an outpatient procedure center via simulation. In: Proceedings of the 2007 Winter Simulation Conference, pp 1489–1497

[143] Iser JH, Denton BT, King RE (2008) Heuristics for balancing operating room and post-anesthesia resources under uncertainty. In: Proceedings of the 2008 Winter Simulation Conference, pp 1601–1608

[144] Jeang A, Chiang AJ (2012) Economic and quality scheduling for effective utilization of operating rooms. Journal of Medical Systems 36:1205–22

[145] Jebali A, Hadj-Alouane A, Ladet P (2003) Performance comparison of two strategies for operating room scheduling. In: International Symposium on Computational Intelligence and Intelligent Informatics

[146] Jebali A, Hadj-Alouane AB, Ladet P (2006) Operating rooms scheduling. International Journal of Production Economics 99:52–62

[147] Jittamai P, Kangwansura T (2011) A hospital admission planning model for emergency and elective patients under stochastic resource requirements and no-shows. In: 2011 IEEE International Conference on Industrial Engineering and Engineering Management, pp 166–170

[148] Joustra PE, de Wit J, Van Dijk NM, Bakker PJM (2011) How to juggle priorities? An interactive tool to provide quantitative support for strategic patient-mix decisions: An ophthalmology case. Health Care Management Science 14:348–360

[149] Jun JB, Jacobson SH, Swisher JR (1999) Application of discrete-event simulation in health care clinics: A survey. The Journal of the Operational Research Society 50:109–123

[150] Keren B, Pliskin J (2011) Optimal timing of joint replacement using mathematical programming and stochastic programming models. Health Care Management Science 14:361–369

[151] Kharraja S, Albert P, Chaabane S (2006) Block scheduling: Toward a master surgical schedule. In: Proceedings of 2006 International Conference on Service Systems and Service Management, pp 429–435

[152] Kim SC, Horowitz I (2002) Scheduling hospital services: The efficacy of elective-surgery quotas. Omega-International Journal of Management Science 30:335–346

[153] Kleinrock L (1964) A delay dependent queue discipline. Naval Research Logistics Quarterly 11:329–341

[154] Kodali BS, Kim D, Bleday R, Flanagan H, Urman RD (2014) Successful strategies for the reduction of operating room turnover times in a tertiary care academic medical center. Journal of Surgical Research 187:403–411

[155] Koenig L, Gu Q (2013) Growth of ambulatory surgical centers, surgery volume, and savings to Medicare. American Journal of Gastroenterology 108:10–15

[156] Kolker A (2009) Process modeling of ICU patient flow: Effect of daily load leveling of elective surgeries on ICU diversion. Journal of Medical Systems 33:27–40

[157] van der Kooij R, Mes M, Hans EW (2014) Simulation framework to analyze operating room release mechanisms. In: Proceedings of the 2014 Winter Simulation Conference, pp 1144–1155

[158] Krempels KH, Panchenko A (2006) An approach for automated surgery scheduling. In: Sixth International Conference on the Practice and Theory of Automated Timetabling

[159] Kuo PC, Schroeder RA, Mahaffey S, Bollinger RR (2003) Optimization of operating room allocation using linear programming techniques. Journal of the American College of Surgeons 197:889–895

[160] Lagergren M (1998) What is the role and contribution of models to management and research in the health services? A view from Europe. European Journal of Operational Research 105:257–266

[161] Lamiri M, Dreo J, Xiaolan X (2007) Operating room planning with random surgery times. In: Proceedings of the 3th IEEE Conference on Automation Science and Engineering, pp 521–6

[162] Lamiri M, Augusto V, Xie X (2008) Patients scheduling in a hospital operating theatre. In: 2008 IEEE International Conference on Automation Science and Engineering, pp 627–632

[163] Lamiri M, Xie X, Dolgui A, Grimaud F (2008) A stochastic model for operating room planning with elective and emergency demand for surgery. European Journal of Operational Research 185:1026–1037

[164] Lamiri M, Xie X, Zhang SG (2008) Column generation approach to operating theater planning with elective and emergency patients. IIE Transactions 40:838–852

[165] Lamiri M, Grimaud F, Xie X (2009) Optimization methods for a stochastic surgery planning problem. International Journal of Production Economics 120:400–410

[166] van der Lans M, Hans EW, Hurink JL, Wullink G, Van Houdenhoven M, Kazemier G (2006) Anticipating urgent surgery in operating room departments, University of Twente

[167] van der Lans M, Hans EW, Hurink JL, Wullink G, Van Houdenhoven M, Kazemier G (2006) Anticipating urgent surgery in operating room departments, University of Twente

[168] Lebowitz P (2003) Schedule the short procedure first to improve or efficiency. AORN Journal 78:651–4

[169] Lee S, Yih Y (2014) Reducing patient-flow delays in surgical suites through determining start-times of surgical cases. European Journal of Operational Research 238:620–629

[170] Lehtonen JM, Torkki P, Peltokorpi A, Moilanen T (2013) Increasing operating room productivity by duration categories and a newsvendor model. International Journal of Health Care Quality Assurance 26:80–92

[171] Leppäniemi A, Jousela I (2014) A traffic-light coding system to organize emergency surgery across surgical disciplines. The British Journal of Surgery 101:134–140

[172] Leslie RJ, Beiko D, Van Vlymen J, Siemens DR (2012) Day of surgery cancellation rates in urology: Identification of modifiable factors. Canadian Urological Association Journal pp 1–8

[173] Lewis HF, Sexton TR, Dolan MA (2011) An efficiency-based multicriteria strategic planning model for ambulatory surgery centers. Journal of Medical Systems 35:1029–1037

[174] Litvak N, van Rijsbergen M, Boucherie RJ, Van Houdenhoven M (2008) Managing the overflow of intensive care patients. European Journal of Operational Research 185:998–1010

[175] Liu Y, Chu C, Wang K (2011) A new heuristic algorithm for the operating room scheduling problem. Computers & Industrial Engineering 61:865–871

[176] Lovejoy WS, Li Y (2002) Hospital operating room capacity expansion. Management Science 48:1369–1387

[177] Luangkesorn KL, Bountourelis T, Schaefer A, Nabors S, Clermont G (2012) The case against utilization: Deceptive performance measures in inpatient care capacity models. In: Proceedings of the 2012 Winter Simulation Conference, p 76

[178] Ma G, Demeulemeester E (2010) Assessing the performance of hospital capacity planning through simulation analysis, KU Leuven

[179] Ma G, Demeulemeester E (2013) A multilevel integrative approach to hospital case mix and capacity planning. Computers & Operations Research 40:2198–2207

[180] Ma G, Beliën J, Demeulemeester E, Wang L (2011) Solving the case mix problem optimally by using branch-and-price algorithms, KU Leuven

[181] Magerlein JM, Martin JB (1978) Surgical demand scheduling: A review. Health Services Research 13:418–433

[182] Mancilla C, Storer RH (2013) Stochastic sequencing of surgeries for a single surgeon operating in parallel operating rooms. IIE Transactions on Healthcare Systems Engineering 3:127–138

[183] Mannino C, Nilssen EJ, Nordlander TE (2012) A pattern based, robust approach to cyclic master surgery scheduling. Journal of Scheduling 15:553–563

[184] Marcon E, Dexter F (2006) Impact of surgical sequencing on post anesthesia care unit staffing. Health Care Management Science 9:87–98

[185] Marcon E, Dexter F (2007) An observational study of surgeons' sequencing of cases and its impact on postanesthesia care unit and holding area staffing requirements at hospitals. Anesthesia and Analgesia 105:119–126

[186] Marcon E, Kharraja S, Simonnet G (2003) The operating theatre planning by the follow-up of the risk of no realization. International Journal of Production Economics 85:83–90

[187] Marcon E, Kharraja S, Smolski N, Luquet B, Viale JP (2003) Determining the number of beds in the postanesthesia care unit: A computer simulation flow approach. Anesthesia and Analgesia 96:1415–1423

[188] Marjamaa RA, Torkki PM, Hirvensalo EJ, Kirvela OA (2009) What is the best workflow for an operating room? A simulation study of five scenarios. Health Care Management Science 12:142–146

[189] Marques I, Captivo ME, Vaz Pato M (2012) An integer programming approach to elective surgery scheduling. OR Spectrum 34:407–427

[190] Marques I, Captivo ME, Vaz Pato M (2014) A bicriteria heuristic for an elective surgery scheduling problem. Health Care Management Science

[191] Marques I, Captivo ME, Vaz Pato M (2014) A bicriteria heuristic for an elective surgery scheduling problem. Health Care Management Science

[192] Marques I, Captivo ME, Vaz Pato M (2014) Scheduling elective surgeries in a Portuguese hospital using a genetic heuristic. Operations Research for Health Care 3:59–72

[193] Masursky D, Dexter F, O'Leary CE, Applegeet C, Nussmeier NA (2008) Long-term forecasting of anesthesia workload in operating rooms from changes in a hospital's local population can be inaccurate. Anesthesia and Analgesia 106:1223–1231

[194] May JH, Spangler WE, Strum DP, Vargas LG (2011) The surgical scheduling problem: Current research and future opportunities. Production and Operations Management 20:392–405

[195] Medpac (2010) Report to congress: Medicare payment policy. Tech. rep., Medicare Payment Advisory Commission

[196] Meskens N, Duvivier D, Lianset A (2013) Multi-objective operating room scheduling considering desiderata of the surgical team. Decision Support Systems 55:650–659

[197] MHallah R, Al-Roomi AH (2014) The planning and scheduling of operating rooms: A simulation approach. Computers & Industrial Engineering 78:235–248

[198] MHallah R, Al-Roomi AH (2014) The planning and scheduling of operating rooms: A simulation approach. Computers & Industrial Engineering 78:235–248

[199] Milliman (2011) 2011 Milliman Medical Index. Tech. rep., Milliman

[200] Min D, Yih Y (2010) An elective surgery scheduling problem considering patient priority. Computers & Operations Research 37:1091–1099

[201] Min D, Yih Y (2010) Scheduling elective surgery under uncertainty and downstream capacity constraints. European Journal of Operational Research 206:642–652

[202] Min D, Yih Y (2014) Managing a patient waiting list with time-dependent priority and adverse events. RAIRO - Operations Research 48:53–74

[203] Molina JM, Framinan JM (2009) Testing planning policies for solving the elective case scheduling phase: A real application. In: 35th International Conference on Operational Research Applied to Health Services

[204] Mulholland MW, Abrahamse P, Bahl V (2005) Linear programming to optimize performance in a department of surgery. Journal of the American College of Surgeons 200:861–868

[205] Niu Q, Peng Q, ElMekkawy T, Tan YY (2007) Performance analysis of the operating room using simulation. In: CDEN and CCEE Conference

[206] Niu Q, Peng Q, Y ElMekkawy T (2013) Improvement in the operating room efficiency using tabu search in simulation. Business Process Management Journal 19:799–818

[207] Niu Q, Peng Q, Y ElMekkawy T (2013) Improvement in the operating room efficiency using tabu search in simulation. Business Process Management Journal 19:799–818

[208] Nouaouri I, Nicolas JC, Jolly D (2009) Scheduling of stabilization surgical cares in case of a disaster. In: 2009 IEEE International Conference on Industrial Engineering and Engineering Management, pp 1974–8

[209] Noyan Ogulata S, Erol R (2003) A hierarchical multiple criteria mathematical programming approach for scheduling general surgery operations in large hospitals. Journal of Medical Systems 27:259–70

[210] Olivares M, Terwiesch C, Cassorla L (2008) Structural estimation of the newsvendor model: An application to reserving operating room time. Management Science 54:41–55

[211] van Oostrum JM, Van Houdenhoven M, Hurink JL, Hans EW, Wullink G, Kazemier G (2008) A master surgical scheduling approach for cyclic scheduling in operating room departments. OR Spectrum 30:355–374

[212] van Oostrum JM, Bredenhoff E, Hans EW (2010) Suitability and managerial implications of a master surgical scheduling approach. Annals of Operations Research 178:91–104

[213] van Oostrum JM, Parlevliet T, Wagelmans APM, Kazemier G (2011) A method for clustering surgical cases to allow master surgical scheduling. INFOR: Information Systems and Operational Research 49:254–260

[214] Ozkarahan I (2000) Allocation of surgeries to operating rooms by goal programing. Journal of Medical Systems 24:339–378

[215] Pandit JJ, Abbott T, Pandit M, Kapila A, Abraham R (2012) Is 'starting on time' useful (or useless) as a surrogate measure for 'surgical theatre efficiency'? Anaesthesia 67:823–32

[216] Paoletti X, Marty J (2007) Consequences of running more operating theatres than anaesthetists to staff them: A stochastic simulation study. British Journal of Anaesthesia 98:462–469

[217] Pariente JMM, Torres JMF, Cia TG (2009) Policies and decision models for solving elective case operating room scheduling. In: International Conference on Computers and Industrial Engineering (CIE 2009), pp 112–117

[218] Paul JA, MacDonald L (2013) Determination of number of dedicated OR's and supporting pricing mechanisms for emergent surgeries. Journal of the Operational Research Society 64:912–924

[219] Persson M, Persson J (2006) Optimization modelling of hospital operating room planning: analyszing strategies and problem settings. In: Annual Conference of OR Applied to Health Services

[220] Persson M, Persson JA (2009) Health economic modeling to support surgery management at a swedish hospital. Omega-International Journal of Management Science 37:853–863

[221] Persson M, Persson JA (2010) Analysing management policies for operating room planning using simulation. Health Care Management Science 13:182–191

[222] Pham DN, Klinkert A (2008) Surgical case scheduling as a generalized job shop scheduling problem. European Journal of Operational Research 185:1011–1025

[223] Pierskalla WP, Brailer DJ (1994) Applications of operations research in health care delivery, Springer-Verlag Berlin Heidelberg, pp 469–505

[224] Pinedo ML (2012) Scheduling: Theory, Algorithms, and Systems. Springer Science & Business Media

[225] Pérez Gladish B, Arenas Parra M, Bilbao Terol A, Rodriguez Uria MV (2005) Management of surgical waiting lists through a possibilistic linear multiobjective programming problem. Applied Mathematics and Computation 167:477–495

[226] Przasnyski ZH (1986) Operating room scheduling: A literature review. AORN Journal 44:67–79

[227] Pulido R, Aguirre AM, Ortega-Mier M, Garcia-Sanchez A, Mendez CA (2014) Managing daily surgery schedules in a teaching hospital: A mixed-integer optimization approach. BMC Health Services Research 14

[228] Pulido R, Aguirre AM, Ortega-Mier M, Garcia-Sanchez A, Mendez CA (2014) Managing daily surgery schedules in a teaching hospital: A mixed-integer optimization approach. BMC Health Services Research 14

[229] Rachuba S, Werners B (2014) A robust approach for scheduling in hospitals using multiple objectives. Journal of the Operational Research Society 65:546–556

[230] Ramis FJ, Palma JL, Baesler FF (2001) The use of simulation for process improvement at an ambulatory surgery center. In: Proceedings of the 2001 Winter Simulation Conference

[231] Riise A, Burke E (2011) Local search for the surgery admission planning problem. Journal of Heuristics 17:389–414

[232] Riise A, Burke E (2011) Local search for the surgery admission planning problem. Journal of Heuristics 17:389–414

[233] Rizk C, Arnaout JP (2012) ACO for the surgical cases assignment problem. Journal of Medical Systems 36:1891–1899

[234] Rizk C, Arnaout JP (2012) ACO for the surgical cases assignment problem. Journal of Medical Systems 36:1891–1899

[235] Roland B, Di Martinelly C, Riane F (2006) Operating theatre optimization: A resource-constrained based solving approach. In: International Conference on Service Systems and Service Management, pp 443–448

[236] Roland B, Di Martinelly C, Riane F, Pochet Y (2010) Scheduling an operating theatre under human resource constraints. Computers & Industrial Engineering 58:212–220

[237] Ruey-Kei C, Yu-Chen Y (2010) Fuzzy-based dynamic scheduling system for health examination. In: 2010 International Conference on Machine Learning and Cybernetics, pp 636–41

[238] Samudra M, Demeulemeester E, Cardoen B (2013) A closer view at the patient surgery planning and scheduling problem: A literature review. Review of Business and Economic Literature (ReBEL) 58:115–140

[239] Samudra M, Demeulemeester E, Cardoen B, Vansteenkiste N, Rademakers FE (2016) Due time driven surgery scheduling. Health Care Management Science pp 1–27

[240] Sandbaek BE, Helgheim BI, Larsen OI, Fasting S (2014) Impact of changed management policies on operating room efficiency. BMC Health Services Research 14

[241] Santibanez P, Begen MA, Atkins D (2007) Surgical block scheduling in a system of hospitals: An application to resource and wait list management in a british columbia health authority. Health Care Management Science 10:269–82

[242] Saremi A, Jula P, ElMekkawy T, Wang GG (2013) Appointment scheduling of outpatient surgical services in a multistage operating room department. International Journal of Production Economics 141:646–658

[243] Sargent R (2005) Verification and validation of simulation models

[244] Schlesinger S, Crosby R, Cagne R, Innis G, Lalwani C, Loch J, Sylvester R, Wright R, Kheir N, Bartos D (1979) Terminology for model credibility. SIMULATION 32:103–104

[245] Schmid V, Doerner KF (2014) Examination and operating room scheduling including optimization of intrahospital routing. Transportation Science 48:59–77

[246] Schoenmeyr T, Dunn PF, Garnarnik D, Levi R, Berger DL, Daily BJ, Levine WC, Sandberg WS (2009) A model for understanding the impacts of demand and capacity on waiting time to enter a congested recovery room. Anesthesiology 110:1293–1304

[247] Sciomachen A, Tanfani E, Testi A (2005) Simulation models for optimal schedules of operating theatres. International Journal of Simulation 6:26–34

[248] Shimazaki H, Shinomoto S (2010) Kernel bandwidth optimization in spike rate estimation. Journal of Computational Neuroscience 29:171–182

[249] Shylo OV, Luangkesorn L, Prokopyev OA, Rajgopal J, Schaefer A (2011) Managing patient backlog in a surgical suite that uses a block-booking scheduling system. In: Proceedings of the 2011 Winter Simulation Conference, pp 1314–1324

[250] Shylo OV, Prokopyev OA, Schaefer AJ (2013) Stochastic operating room scheduling for high-volume specialties under block booking. INFORMS Journal on Computing 25:682–692

[251] Siciliani LMB, Moran V (2013) Waiting Time Policies in the Health Sector. OECD Publishing

[252] Sieber TJ, Leibundgut DL (2002) Operating room management and strategies in Switzerland: Results of a survey. European Journal of Anaesthesiology 19:415–423

[253] Slack N (1999) The Blackwell Encyclopedic Dictionary of Operations Management. Wiley

[254] Smith-Daniels VL, Schweikhart SB, Smith-Daniels DE (1988) Capacity management in health care services: Review and future research directions. Decision Sciences 19:889–919

[255] Sobolev BG, Sanchez V, Vasilakis C (2011) Systematic review of the use of computer simulation modeling of patient flow in surgical care. Journal of Medical Systems 35:1–16

[256] Souki M, Rebai A (2009) Memetic differential evolution algorithm for operating room scheduling. In: International Conference on Computers and Industrial Engineering (CIE 2009), pp 845–850

[257] Souki M, Ben Youssef S, Rebai A (2009) Memetic algorithm for operating room admissions. In: International Conference on Computers and Industrial Engineering (CIE 2009), pp 519–524

[258] Sperandio F, Gomes C, Borges J, Carvalho Brito A, Almada-Lobo B (2014) An intelligent decision support system for the operating theater: A case study. IEEE Transactions on Automation Science and Engineering 11:265–273

[259] Stanciu A, Vargas LG, May JH (2010) A revenue management approach for managing operating room capacity. In: Proceedings of the 2010 Winter Simulation Conference, pp 2444–54

[260] Stanford D, Taylor P, Ziedins I (2014) Waiting time distributions in the accumulating priority queue. Queueing Systems 77:297–330

[261] Stanford D, Taylor P, Ziedins I (2014) Waiting time distributions in the accumulating priority queue. Queueing Systems 77:297–330

[262] Steins K, Persson F, Holmer M (2010) Increasing utilization in a hospital operating department using simulation modeling. Simulation 86:463–480

[263] Strum DP, May JH, Vargas LG (2000) Modeling the uncertainty of surgical procedure times - comparison of log-normal and normal models. Anesthesiology 92:1160–1167

[264] Stuart K, Kozan E (2012) Reactive scheduling model for the operating theatre. Flexible Services and Manufacturing Journal 24:400–421

[265] Tan Y, El Mekkawy T, Peng Q, Oppenheimer L (2007) Mathematical programming for the scheduling of elective patients in the operating room department. In: Proceedings of the Canadian Engineering Education Association

[266] Tancrez JS, Roland B, Cordier JP, Riane F (2009) How stochasticity and emergencies disrupt the surgical schedule, vol 189, Springer Berlin / Heidelberg, pp 221–239

[267] Tancrez JS, Roland B, Cordier JP, Riane F (2013) Assessing the impact of stochasticity for operating theater sizing. Decision Support Systems 55:616–628

[268] Tanfani E, Testi A (2010) Improving surgery department performance via simulation and optimization. In: 2010 IEEE Workshop on Health Care Management (WHCM), p 6 pp.

[269] Tanfani E, Testi A (2010) A pre-assignment heuristic algorithm for the master surgical schedule problem (mssp). Annals of Operations Research 178:105–119

[270] Testi A, Tanfani E (2009) Tactical and operational decisions for operating room planning: Efficiency and welfare implications. Health Care Management Science 12:363–373

[271] Testi A, Tanfani E, Torre G (2007) A three-phase approach for operating theatre schedules. Health Care Management Science 10:163–72

[272] Testi A, Tanfani E, Valente R, Ansaldo G, Torre G (2008) Prioritizing surgical waiting lists. Journal of Evaluation in Clinical Practice 14:59 – 64

[273] Tewari A, Giering MJ, Raghunathan A (2011) Parametric characterization of multimodal distributions with non-gaussian modes. In: 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW), pp 286–292

[274] Trivedi P, Zimmer D (2006) Copula modeling: An introduction for practitioners. Foundations and Trends® in Econometrics 1:1–111

[275] Tsoy G, Arnaout JP, Smith T, Rabadi G (2004) A genetic algorithm approach for surgery operating rooms scheduling problem. In: 25th National Conference of the American Society for Engineering Management, pp 299–304

[276] Tyler DC, Pasquariello CA, Chen CH (2003) Determining optimum operating room utilization. Anesthesia and Analgesia 96:1114–1121

[277] Utley M, Worthington D (2012) Capacity planning, vol 168, Springer US, pp 11–30

[278] Van Houdenhoven M, Hans EW, Klein J, Wullink G, Kazemier G (2007) A norm utilisation for scarce hospital resources: Evidence from operating rooms in a Dutch university hospital. Journal of Medical Systems 31:231–236

[279] Van Houdenhoven M, van Oostrum JM, Hans EW, Wullink G, Kazemier G (2007) Improving operating room efficiency by applying bin-packing and portfolio techniques to surgical case scheduling. Anesthesia and Analgesia 105:707–714

[280] Van Houdenhoven M, van Oostrum JM, Wullink G, Hans EW, Hurink JL, Bakker J, Kazemier G (2008) Fewer intensive care unit refusals and a higher capacity utilization by using a cyclic surgical case schedule. Journal of Critical Care 23:222–226

[281] Van Huele C, Vanhoucke M (2014) Analysis of the integration of the physician rostering problem and the surgery scheduling problem. Journal of Medical Systems 38

[282] Vanberkel PT, Blake JT (2007) A comprehensive simulation for wait time reduction and capacity planning applied in general surgery. Health Care Management Science 10:373–85

[283] Vanberkel PT, Boucherie RJ, Hans EW, Hurink JL, Litvak N (2009) A survey of health care models that encompass multiple departments, University of Twente

[284] Vanberkel PT, Boucherie RJ, Hans EW, Hurink JL, van Lent WAM, van Harten WH (2011) Accounting for inpatient wards when developing master surgical schedules. Anesthesia and Analgesia 112:1472–1479

[285] Vanberkel PT, Boucherie RJ, Hans EW, Hurink JL, van Lent WAM, van Harten WH (2011) An exact approach for relating recovering surgical patient workload to the master surgical schedule. Journal of the Operational Research Society 62:1851–1860

[286] Vanberkel PT, Boucherie RJ, Hans EW, Hurink JL (2014) Optimizing the strategic patient mix combining queueing theory and dynamic programming. Computers & Operations Research 43:271–279

[287] Vansteenkiste N, Lamote C, Vandersmissen J, Luysmans P, Monnens P, De Voldere G, Kips J, Rademakers FE (2012) Reallocation of operating room capacity using the due-time model. Medical Care 50:779–784

[288] Vansteenkiste N, Lamote C, Vandersmissen J, Luysmans P, Monnens P, De Voldere G, Kips J, Rademakers FE (2012) Reallocation of operating room capacity using the due-time model. Medical Care 50:779–784

[289] van Veen-Berkx E, Bitter J, Elkhuizen SG, Buhre WF, Kalkman CJ, Gooszen HG, Kazemier G (2014) The influence of anesthesia-controlled time on operating room scheduling in dutch university medical centres. Can J Anaesth 61:524–32

[290] Velasquez R, Melo T, Kufer KH (2008) Tactical operating theatre scheduling: Efficient appointment assignment. In: Operations Research Proceedings 2007, pp 303–308

[291] Viberg N, Molin R, Forsberg B (2011) Swedish waiting times for health care in an international perspective. Tech. rep., sALAR

[292] Vijayakumar B, Parikh PJ, Scott R, Barnes A, Gallimore J (2013) A dual bin-packing approach to scheduling surgical cases at a publicly-funded hospital. European Journal of Operational Research 224:583–591

[293] Vijayakumar B, Parikh PJ, Scott R, Barnes A, Gallimore J (2013) A dual bin-packing approach to scheduling surgical cases at a publicly-funded hospital. European Journal of Operational Research 224:583–591

[294] Vissers J, Bertrand J, De Vries G (2001) A framework for production control in health care organizations. Production Planning & Control 12:591–604

[295] Vissers J, Adan I, Bekkers J (2005) Patient mix optimization in tactical cardiothoracic surgery planning: A case study. IMA Journal of Management Mathematics 16:281–304,304

[296] Wachtel RE, Dexter F (2007) A simple method for deciding when patients should be ready on the day of surgery without procedure-specific data. Anesthesia and Analgesia 105:127–140

[297] Wachtel RE, Dexter F (2008) Tactical increases in operating room block time for capacity planning should not be based on utilization. Anesthesia and Analgesia 106:215–26

[298] Wachtel RE, Dexter F (2009) Influence of the operating room schedule on tardiness from scheduled start times. Anesthesia and Analgesia 108:1889–1901

[299] Wachtel RE, Dexter F (2009) Reducing tardiness from scheduled start times by making adjustments to the operating room schedule. Anesthesia and Analgesia 108:1902–1909

[300] Wang D, Xu JP (2008) A fuzzy multi-objective optimizing scheduling for operation room in hospital. In: 2008 IEEE International Conference on Industrial Engineering and Engineering Management, pp 614–618

[301] Wang QN (2004) Modeling and analysis of high risk patient queues. European Journal of Operational Research 155:502–515

[302] Wang T, Meskens N, Duvivier D (2012) A comparison of mixed-integer programming and constraint programming models for scheduling problem in operating theatres. In: 2012 International Conference on Information Systems, Logistics and Supply Chain

[303] Wang Y, Tang J, Fung RYK (2014) A column-generation-based heuristic algorithm for solving operating theater planning problem under stochastic demand and surgery cancellation risk. International Journal of Production Economics 158:28–36

[304] Wullink G, Van Houdenhoven M, Hans EW, van Oostrum JM, van der Lans M, Kazemier G (2007) Closing emergency operating rooms improves efficiency. Journal of Medical Systems 31:543–546

[305] Xiang W, Yin J, Lim G (2013) Modified ant colony algorithm for surgery scheduling under multiresource constraints. Advances in Information Sciences and Service Sciences 5:810

[306] Xiang W, Yin J, Lim G (2014) A short-term operating room surgery scheduling problem integrating multiple nurses roster constraints. Artificial Intelligence in Medicine

[307] Xue W, Yan Z, Barnett R, Fleisher L, Liu R (2013) Dynamics of elective case cancellation for inpatient and outpatient in an academic center. Journal of Anesthesia & Clinical Research 4:314

[308] Ya L, Chengbin C, Kanliang W (2010) Aggregated state dynamic programming for operating theater planning. In: 2010 IEEE International Conference on Automation Science and Engineering, pp 1013–18

[309] Yu W, Jiafu T, Gang Q (2010) A genetic algorithm for solving patient-priority-based elective surgery scheduling problem. In: Life System Modeling and Intelligent Computing. International Conference on Life System Modeling and Simulation, LSMS 2010, and International Conference on Intelligent Computing for Sustainable Energy and Environment, ICSEE 2010, pp 297–304

[310] Zhang B, Murali P, Dessouky MM, Belson D (2009) A mixed integer programming approach for allocating operating room capacity. Journal of the Operational Research Society 60:663–673

[311] Zheng Z, Xiaolan X, Na G (2012) Promise surgery start times and implementation strategies. In: 2012 IEEE International Conference on Automation Science and Engineering, pp 143–149

[312] Zheng Z, Xiaolan X, Na G (2014) Dynamic surgery assignment of multiple operating rooms with planned surgeon arrival times. IEEE Transactions on Automation Science and Engineering 11:680–691

[313] Zheng Z, Xiaolan X, Na G (2014) Simulation-based surgery appointment sequencing and scheduling of multiple operating rooms. In: 2014 IEEE International Conference on Automation Science and Engineering, pp 399–404

[314] Zonderland ME, Boucherie RJ, Litvak N, Vleggeert-Lankamp C (2010) Planning and scheduling of semi-urgent surgeries. Health Care Management Science 13:256–267

Doctoral Dissertations from the Faculty of Business and Economics

Doctoral dissertations from the Faculty of Economics and Business, see: http://www.kuleuven.ac.be/doctoraatsverdediging/archief.htm.